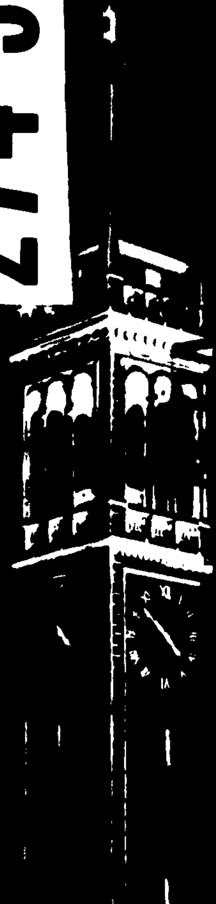


On the Stability of Randomly Varying Systems

by

B. H. Bharucha

Series No. 60, Issue No. 381
July 5, 1961
AF 19(604)-5466

ELECTRONICS RESEARCH LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY CALIFORNIA

AFCRL-62-38

Electronics Research Laboratory
University of California
Berkeley, California

ON THE STABILITY OF RANDOMLY VARYING SYSTEMS

by

B. H. Bharucha

Institute of Engineering Research
Series No. 60, Issue No. 381

Contract No. AF 19(604)-5466
Project No. 5632
Task No. 56320

Scientific Report No. 7

July 5, 1961

Prepared for
Electronics Research Directorate
Air Force Cambridge Research Laboratories
Office of Aerospace Research
United States Air Force
Bedford, Massachusetts

ACKNOWLEDGMENT

The author wishes to thank Professor A. R. Bergen for acting as Research Advisor during the course of this investigation, and Mr. Barry Whalen for the many stimulating discussions that he has had with him.

ABSTRACT

This study is concerned with the stability of random systems, that is, systems whose internal characteristics are governed by probability laws.

Concepts of stability appropriate to random differential systems are formulated and discussed. Precise definitions of stability are stated and theorems interrelating these definitions are proven. Some particular types of stability investigated are, in the mean norm, in the ith moment, in probability, almost sure, and almost uniform-in- ω.

Particular attention is focused on the random linear (vector) differential equation with piecewise constant parameters:

$$\dot{x} = A_k x, \quad t_{k-1} \leq t < t_k, \quad k = 1, 2, \ldots$$

Two statistical structures are studied:

(a) $\{A_k (t_k - t_{k-1})\}$ is an independent, identically distributed, random process;

(b) $\{A_k(t_k - t_{k-1})\}$ is a finite Markov chain.

The ith moment of the solutions is obtained explicitly and is studied with respect to asymptotic behavior, and with respect to the set structure of the Markov chain in case (b). Sufficient conditions for exponential asymptotic stability in the ith moment are derived. (In case (a), these conditions are also necessary.) It is shown that if i is even, then these sufficient conditions also imply almost sure asymptotic stability.

For random linear systems, it is shown that if the trivial solution of the homogeneous equation is exponentially asymptotically stable in the mean norm, then certain stochastic equivalents of "bounded output to every bounded input" are true.

With minor modifications, all of the above results are equally applicable to random difference equations.

TABLE OF CONTENTS

I. INTRODUCTION AND SUMMARY

1.1 General Introduction

By a random system one understands here a system some or all of whose internal characteristics are governed by probability laws. More precisely, it is assumed that the unforced system can be represented by an ensemble of equations indexed by the element ω of a probability space. Thus, for example, the system may have a representation as a linear differential equation with random processes as coefficients. This work is concerned with the stability of systems described by random differential or difference equations. The problems of measurement, identification and optimization are not touched upon.

A variety of physical situations can be represented by random differential equations. Two examples in the field of systems are: (i) an adaptive system which compensates for effects of an external random disturbance; and (ii) a control system whose parameters are undergoing noise modulation. An illustration of a random parameter is the gain of the control surface of an aircraft in flight through a medium, possibly turbulent, of rapidly changing and incompletely known characteristics. In the field of circuits, lumped networks containing randomly varying elements can be described by random differential equations; likewise, long transmission lines with characteristics that are nonuniform along the length of the line.[1] Finally, problems of wave propagation and scattering in a random medium also lend themselves to a random differential equation representation.[2]

Random difference equations arise in most of the above situations if the coefficients of the equations are random but piecewise constant in the independent variable.[3,4] They occur also in the field of quantum mechanics in the study of energy levels in random lattices.[5] A final example is the area of randomly sampled systems, that is, sampled-data systems whose sampling intervals are random variables.[6] The random sampling may result, for instance, from economic considerations, as in the time sharing of a digital computer that controls several processes; from requirements of secrecy and anti-jam protection, as in military communication systems; from unavoidable perturbations on the nominal period of a "periodically" sampled system. Finally, there are situations which though not involving random sampling per se, can be formulated in terms of a sampler which "skips" or fails to operate in some random manner. Two such examples are data transmission links and scan radars, where the received data may be rejected at some time instants because of excessive noise.

1.2 Some Previous Work on the Stability of Random Systems

For random linear systems with continuous parameter variations, some results are available. For first-order systems Rosenbloom[7] has expressed the output moments in terms of the characteristic function of the indefinite time integral of the parameter process; he shows in particular that if the parameter process is stationary and Gaussian, then for a step input the first and second output moments may become unbounded, whereas the output approaches one in probability, as $t \to \infty$. Tikhonov[8] has considered the first-order case where the input and parameter processes are jointly Gaussian. The random linear differential system containing one purely random coefficient has been studied in some detail by Samuels and Eringen.[1] Samuels[9] has further developed the theory and extended it to systems containing one narrow-band random parameter. He has considered also linear systems with dependent parameter processes and an independent input and has arrived at specific results in some special cases. Bergen[10] has also studied the linear differential system containing only one purely random coefficient, and has found necessary and sufficient conditions for the mean square error to remain bounded when the system is excited by an independent input. Zadeh[11,12] has investigated a very general class of random linear systems admitting of a certain integral representation, and has demonstrated an integral relation between the output covariance, and the input and "system" covariance, similar to that for deterministic linear systems.

Bertram and Sarachik[13] have extended "Lyapunov's Second Method" to random systems. By working with the expectation of the total time derivative of the Lyapunov function along the system trajectories and measuring stability in the sense of the mean norm, they have arrived at theorems analogous to those in the deterministic case.

For discrete random linear systems (and continuous systems with piecewise constant parameters), the problem consists of studying the behavior of a given initial state vector undergoing a succession of random linear transformations. Kalman[6] derived the necessary and sufficient conditions for mean square stability of an nth-order randomly sampled system whose sampling intervals form a sequence of independent, identically distributed, random variables. For the random linear nth-order system with piecewise constant parameters:
$\dot{x} = A_k x$, $t_{k-1} \leq t < t_k$, $k = 1, 2, \ldots$, Bergen[3] has found necessary and sufficient conditions for asymptotic stability in the second moment when $\{A_k\}$ is a deterministic sequence of the form $\{B, C, B, C, \ldots\}$ and $\{t_k - t_{k-1}\}$ is an independent, identically distributed, random sequence. This author[4] has investigated the asymptotic stability in the ith moment for the case where $\{A_k(t_k - t_{k-1})\}$ is a finite Markov chain (see Chaps. III, IV). All three above authors use the device of the Kronecker product of matrices, first used by Bellman[14] in studying the asymptotic behavior of products of independent random linear transformations. In a recent paper on products of random matrices, Furstenberg and Kesten[15] have found general conditions for the convergence of the random sequence $\{n^{-1} \log \| X_n \ldots X_2 X_1 \| \}$ with $\{X_n\}$ a stationary stochastic process with values in the set of k x k matrices, and have deduced the asymptotic normality of $\log(X_n X_{n-1} \ldots X_1)$, a result earlier conjectured by Bellman.[14]

1.3 Outline of Present Study

In general, the problem of stability is to determine the extent to which a set of properties of a system remains invariant under a specified set of disturbances or alterations on the system, on the initial conditions, etc. For deterministic systems, a common stability problem is the study of deviations of the system state vector from a given equilibrium state (or equivalently from a given solution) when initial conditions are close to this equilibrium state; the various stability concepts are generated by imposing requirements that the magnitude of the deviations remain small for all time, that it returns to zero as $t \to \infty$, that it satisfies uniformity conditions with respect to initial conditions, etc. The corresponding problem for random systems is considered here, the deviation now being measured in some stochastic sense, e.g., in probability, almost surely (with probability one), almost uniformly (-in- ω), in the ith mean, etc.

In Chap. II, concepts of stability appropriate to random differential systems are formulated and discussed. Precise definitions of stability are given and theorems interrelating these definitions are proven. In the beginning of the chapter, numerous definitions of stability are stated to provide the motivation for the definitions in the random case. Although here, and in the remaining chapters, the system is assumed to be a differential one, by making small changes it is possible to apply the material presented to other types of systems, e.g., to systems of random difference equations.

Chapter III is devoted to the stability analysis of the random linear differential system with piecewise constant coefficients:

$$\dot{x} = A_k x, \quad t_{k-1} \leq t < t_k, \quad k = 1, 2, \ldots,$$

where $x(t)$ is an n-vector and A_k is a constant n x n matrix. The behavior of the system is completely determined by the random sequence $\{A_k(t_k - t_{k-1})\}$, assumed in Chap. III to be an independent, identically distributed, random process. Some particular results obtained are the necessary and sufficient conditions for asymptotic stability in the ith moment and sufficient conditions for almost sure asymptotic stability.

Chapter IV considers the same system as Chap. III except that $\{A_k(t_k-t_{k-1})\}$ is now a finite Markov chain. The results obtained are almost identical to those of Chap. III. In fact, all the results of Chap. III can be deduced from those of Chap. IV (and Appendix E). The independent case is treated separately in Chap. III because it seems preferable to do so from an expository point of view.

The preceding chapters have considered the unforced system. In practice, however, it is the behavior of the system in the presence of an input that is of interest. In Chap. V, a theorem is proven for random linear systems which shows that, if the unforced system is exponentially asymptotically stable in the mean norm, then some stochastic equivalents of "bounded output to every bounded input" are true.

II. STABILITY OF RANDOM SYSTEMS

Summary

Concepts of stability appropriate to random differential systems are formulated and discussed. Precise definitions of stability are given and theorems interrelating these definitions are proven. In the beginning of the chapter numerous definitions of stability of deterministic differential systems are stated to provide the motivation for the definitions in the random case. Although the system is assumed to be a differential one, by making small changes it is possible to apply the material presented here to other types of systems, e.g., systems of random difference equations.

2.1 Deterministic Systems

Consider the differential equation

$$\dot{x} = f(x, t) \tag{2.1}$$

where x is an element of the real, (normed) vector n-space R^n, and f is defined on $R^n X \{t:t \geq 0\}$. The function f is assumed to be sufficiently smooth so that solutions of (2.1) exist and are unique for all $t \geq 0$ for all initial values $x_0 = x(t_0)$; $t_0 \geq 0$, and are continuous in (x_0, t_0).[35] $g(t, x_0, t_0)$ will be used to designate that solution which satisfies the initial condition $g(t_0, x_0, t_0) = x_0$. By an abuse of notation, $x(t)$ also will be used to designate a solution but the context will be such as to preclude any ambiguity with the variable x of the differential equation (2.1).

Without loss of generality it is assumed that $f(0, t) = 0$ for all $t \geq 0$ so that $x \equiv 0$ is always a solution, called the trivial solution. For, given any solution of a differential system, it is always possible by introducing a change of variables in the differential equations to arrive at a new differential system[16] which satisfies the relation $f(0, t) = 0$ for all $t \geq 0$.

The term stability as considered here is the study of the deviation, from the trivial solution, of solutions corresponding to nonzero initial conditions.

DEFINITIONS *: The trivial solution $g(t, 0, t_0) = 0$ is called

(i) <u>stable</u> if given $\epsilon > 0$, t_0, there exists $\delta(\epsilon, t_0) > 0$ such that $\|x_0\| < \delta$ implies $\|g(t, x_0, t_0)\| < \epsilon$ for all $t \geq t_0$.

(ii) <u>uniformly stable</u> if given $\epsilon > 0$, there exists $\delta(\epsilon) > 0$ such that for all t_0, $\|x_0\| < \delta$ implies $\|g(t, x_0, t_0)\| < \epsilon$ for all $t \geq t_0$.

(iii) <u>quasi-asymptotically stable</u> if given t_0, there exists $\delta(t_0) > 0$ such that $\|x_0\| < \delta$ implies $g(t, x_0, t_0) \to 0$ as $t \to \infty$.

(iv) <u>asymptotically stable</u> if it is stable and quasi-asymptotically stable.

(v) <u>quasi-equiasymptotically</u> stable if given any t_0, there exists $\delta(t_0) > 0$ such that $\|x_0\| < \delta$ implies $g(t, x_0, t_0) \to 0$ as $t \to \infty$ uniformly on $\|x_0\| < \delta$.

---

* It is understood that $t \geq t_0 \geq 0$.

(vi) equiasymptotically stable if it is stable and quasi-equiasymptotically stable.

(vii) quasi-uniformly asymptotically stable if there exists $\delta > 0$ such that for all t_0, $\|x_0\| < \delta$ implies $g(t, x_0, t_0) \longrightarrow 0$ as $t \longrightarrow \infty$ uniformly on $t_0 \geq 0$, $\|x_0\| < \delta$.

(viii) uniformly asymptotically stable if it is uniformly stable and quasi-uniformly asymptotically stable.

(ix) exponentially asymptotically stable if there exists a $\nu > 0$, and given any $\epsilon > 0$, there is a corresponding $\delta(\epsilon) > 0$ such that for all t_0, $\|x_0\| < \delta$ implies $\|g(t, x_0, t_0)\| < \epsilon \exp[-\nu(t-t_0)]$ for all $t \geq t_0$.

Implications such as the following ones are obvious:

(ii) $\Longrightarrow$ (i)

(vii) $\Longrightarrow$ (v) $\Longrightarrow$ (iii)

(viii) $\Longrightarrow$ (i) - (vii)

For other implications, examples, and discussion of the definitions, see Antosiewicz,[17] Massera,[18] Kalman and Bertram.[19] The latter two references also consider concepts of stability in-the-large.

If the system is linear, that is if $f(x, t)$ is a linear in x, then (viii) and (ix) are equivalent; in fact, for (ix) there is a constant $K \geq 1$ such that it is possible to choose $\delta(\epsilon) = \epsilon / K$ so that $\|g(t, x_0, t_0)\| < K\|x_0\| \exp[-\nu(t-t_0)]$ for all $t \geq t_0$. See Massera.[18]

Definition (i) requires that the trivial solution be stable for every initial time t_0. Since the solution functions are assumed to be continuous in the initial value x_0 it suffices to have the trivial solution stable for some t_0; namely the trivial solution is stable if (and only if) there exists some number t_0 with the property that given any $\epsilon > 0$ there is a $\delta(\epsilon, t_0) > 0$ such that $\|x_0\| < \delta$ implies $\|g(t, x_0, t_0)\| < \epsilon$ for all $t \geq t_0$. See Kalman and Bertram.[19]

Similarly for definition (iii), it suffices to have quasi-asymptotic stability for some initial time t_0; that is in order that the trivial solution be quasi-asymptotically stable, it suffices to have some t_0 and a corresponding $\delta(t_0) > 0$ such that $\| x_0 \| < \delta$ implies $g(t, x_0, t_0) \longrightarrow 0$ as $t \longrightarrow \infty$. For, given any other initial time t_0', consider the mapping of the set $E_0 = \{x_0 : \| x_0 \| < \delta(t_0)\}$ under the solution function to the set $E_0' = \{x_0' = g(t_0', x_0, t_0) : \| x_0 \| < \delta(t_0)\}$. Since the solutions are unique and continuous in the initial state, the mapping $E_0 \longrightarrow E_0'$ is a homeomorphism. Further, E_0' contains the point $x_0' = 0$ for this is the map of the point $x_0 = 0$. Hence, there is a set $E_{\delta'} = \{x_0' : x_0' < \delta'\} \subset E_0'$ for some $\delta' > 0$, and $g(t_0, x_0', t_0') \in E_0$ for every $x_0' \in E_{\delta'}$, which implies that $g(t, x_0', t_0') \longrightarrow 0$ as $t \longrightarrow \infty$.

All of the above stability properties are independent of the choice of the norm. For, given any two norms $\| \cdot \|^{(1)}$ and $\| \cdot \|^{(2)}$ defined for the elements x of a finite-dimensional vector space, there exist two constants k_{12}, k_{21}, such that (see Householder[20], Naimark[36]).

$$k_{21} \| x \|^{(2)} \leq \| x \|^{(1)} \leq k_{12} \| x \|^{(2)} \qquad \text{for all } x.$$

2.2 Random Systems

The differential equation (2.1) now becomes

$$\dot{x} = f(x, t; \omega) \tag{2.2}$$

where ω is a point in a probability space Ω. For every fixed ω, f satisfies the assumptions made at the beginning of the previous section on deterministic systems. The time interval of definition of f, $[0, \infty)$, is fixed, i.e., does not depend on the choice of ω. Also, $f(0, t; \omega) = 0$ on $[0, \infty) \times \Omega$ so that again $x \equiv 0$ is always a solution.

To every (x_0, t_0) there corresponds the set of solutions $\{g(t, x_0, t_0; \omega), \omega \in \Omega\}$ with $g(t_0, x_0, t_0; \omega) = x_0$ for all ω; that is, to every (x_0, t_0) there corresponds a random process, namely, a family of random variables indexed by t. The problem of stability is that of studying the probabilistic departure, from the trivial solution, of the class of random processes generated when x_0 is taken to be different from zero. The definitions given in the previous section are to be modified so as to take into account the dependence on ω of the solution function $g(t, x_0, t_0; \omega)$. This dependence on ω, however, will not be exhibited always by the notation -- in most cases solutions will be designated by $g(t, x_0, t_0)$.

Thus from the previous definition of stable, one may obtain

DEFINITION (i-a): The trivial solution $g(t, 0, t_0) = 0$ is called stable in the mean norm if given $\epsilon > 0$, t_0, there exists $\delta(\epsilon, t_0) > 0$ such that $\| x_0 \| < \delta$ implies $E \| g(t, x_0, t_0) \| < \epsilon$ for all $t \geq t_0$.

In connection with the above definition it is useful to note that

(1) Norms such as

$$\| x \|_p = (\sum | x_i |^p)^{1/p} \quad , \quad p \geq 1 \tag{2. 3}$$

may not be desirable from the viewpoint of analysis for they involve the operations of taking the absolute value and extracting the pth root. However, for any scalar random variable Z,[21]

$$(E | Z |^{r_1})^{1/r_1} \leq (E | Z |^{r_2})^{1/r_2} \quad , \quad 0 \leq r_1 < r_2 \tag{2. 4}$$

Then for any fixed $p \geq 1$, it follows, by setting $Z = | x_i |^p$, $r_1 = 1/p$, and $r_2 = 1$, that

$$(E \| x \|_p)^p = \left\{ E(\sum | x_i |^p)^{1/p} \right\}^p \leq E \sum | x_i |^p = E \| x \|_p^p \tag{2. 5}$$

Hence, in establishing the property of stable in the mean norm with respect to the norm $\| x \|_p$, it suffices to show that $E \; | g_i |^p < \epsilon$ rather than $E \| g \|_p < \epsilon$ as in definition (i). The quantity $\| x \|_p^p$ is not endowed with the homogeneity property of the norm, $\| ax \| = | a | \; \| x \|$, on which rests the invariance property:

(2) As in the deterministic case, the choice of the particular norm is irrelevant -- the property of stability in the mean norm is invariant under changes of the norm.

It follows from (1) and (2) that stability in the mean norm is a consequence of mean square stability, defined as:

DEFINITION (i-b):* The trivial solution is called <u>mean square stable</u> if to every $\epsilon > 0$, t_0, there corresponds a $\delta(\epsilon, t_0) > 0$ such that $\|x_0\| < \delta$ implies

$$E \sum_{i=1}^{n} \left(g_i(t, x_0, t_0)\right)^2 < \epsilon \quad \text{for all } t \geq t_0.$$

KRONECKER AND POWER PRODUCTS. The theory of Kronecker and power products of matrices is connected with the problem of stability in the mean square sense, or more generally, in the sense of the ith moment, of linear systems. Appendix A defines the Kronecker and the power products and states some of their properties. Stability concepts based on the Kronecker and the power product are motivated and developed in Chaps III and IV. The brief discussion below is included here for the sake of completeness.

The Kronecker product of the n-vector $x = (\xi_1, \xi_2, \ldots \xi_n)$ by the m-vector $y = (\eta_1, \eta_2, \ldots, \eta_m)$ is the nm-vector

$$x \otimes y = (\xi_1\eta_1, \xi_1\eta_2, \ldots, \xi_1\eta_m, \xi_2\eta_1, \xi_2\eta_2, \ldots, \xi_n\eta_m)$$

having as components all possible products of a component of x by a component of y. The Kronecker product of x by itself is the n-vector $x \otimes x$ and is denoted by $x_{[2]}$. The ith self-Kronecker product may be defined by the recursive relation:

$$x_{[i]} = x \otimes x_{[i-1]}, \quad i = 2, 3, \ldots$$

($= x_{[i-1]} \otimes x$ since the product is associative). The vector $x_{[1]}$ is taken to be identically x .

The power product of x by itself is the $n(n+1)/2$ vector

$$x_{(2)} = (\xi_1\xi_1, \xi_1\xi_2, \ldots, \xi_1\xi_n, \xi_2\xi_2, \xi_2\xi_3, \ldots, \xi_n\xi_n)$$

---

* For an nth-order scalar differential equation, the term mean square stable is used often in the sense that the expectation of the square of the (scalar) solution function remains small. See Refs. 9 and 10.

having as components distinct products of components of x. A higher power $x_{(i)}$ is similarly defined as the vector having as components the $\binom{n+i-1}{i}$ distinct ith-order products of components of x. The lexicographic ordering is a possible ordering for the components of $x_{(i)}$, and is the one obtained from the Kronecker product $x_{[i]}$ by inspecting successively the components of $x_{[i]}$ starting with the first component and deleting any component which is identical to any preceding component. Thus, if $x = (u, v)$, then $x_{[2]} = (u^2, uv, vu, v^2)$, and $x_{(2)} = (u^2, uv, v^2)$.

Returning to the discussion of stability,

DEFINITION (i-c): The trivial solution is said to be <u>stable in the ith moment</u> if given any $\epsilon > 0$, t_0, there exists $\delta(\epsilon, t_0) > 0$ such that $\|x_0\| < \delta$ implies $\|Eg_{[i]}(t, x_0, t_0)\| < \epsilon$ for all $t \geq t_0$.

The above definition remains unchanged if the Kronecker product $g_{[i]}$ is replaced by the power product $g_{(2)}$. Such a replacement will not, in general, alter in any significant way most of the statements made in this report. It turns out that for theoretical analyses it is simpler to use the Kronecker product for one is not plagued by considerations of ordering of the components of the vectors and of the matrices. Hence, the Kronecker product will be used exclusively in the sequel. However, when numerical computations have to be performed in a particular case, the problem can be formulated in terms of the power product for the smaller order of the power product vector can be of distinct computational advantage.

It is apparent from the definition of the Kronecker product and from definition (i-c) that stability in the ith moment requires not only that the ith moments of the individual components of g be small, but also that all ith-degree cross moments of components of g be small. The value of this latter requirement on the cross moments becomes clear in Chaps III and IV. But it should be mentioned here that if i is even, then the requirement on the cross moments is, in a sense, superfluous. This follows from the inequality

$$E\prod_j |Z_j^{i_j}| \le \prod_j E|Z_j^i|^{i_j/i} \tag{2.6}$$

$$\sum_j i_j = i, \quad i_j \ge 0, \quad j = 1, 2, \ldots, n$$

where the Z_j are scalar random variables. The inequality is a modified form of the Hölder inequality and is proven in Appendix B. Hence, for i even, if for the vector $g = (g_1, g_2, \ldots, g_n)$, Eg_j^i is small for all j, then so is $E(g_1^{i_1} g_2^{i_2} \ldots g_n^{i_n})$, $\sum_{j=1}^{n} i_j = i$. Consequently, definition (i-c) of stable in the ith moment for even i remains unchanged if $\| Eg_{[i]} \|$ is replaced by $E \sum_{j=1}^{n} g_j^i$. In particular, stability in the second moment, (i-c), is equivalent to mean square stability, (i-b). Moreover, by inequality (2.4), it follows that if i is even then stability in the ith moment implies stability in the jth moment for all even $j \le i$. Hence, it also follows that stability in the ith moment for any even i implies stability in the mean norm (irrespective of the norm). The various properties established above can be of value in the stability analysis of a given system where the individual nature of the problem may make it more pertinent or expedient to consider a particular type, rather than some other type, of stability.

Unlike the three preceding definitions, the two following definitions do not involve any "averaging" operations.

DEFINITION: The trivial solution is called

(i-d) stable in probability if given $\epsilon > 0$, $\eta > 0$, t_0, there exists $\delta(\epsilon, \eta, t_0) > 0$ such that $\| x_0 \| < \delta$ implies $P[\, \| g(t, x_0, t_0) \| < \epsilon \,] > 1 - \eta$ for all $t \ge t_0$.

(i-e) stable almost surely if there exists an ω-set S with $P(S) = 1$, and given $\omega \epsilon S$, $\epsilon > 0$, t_0, there exists $\delta_\omega(\epsilon, t_0) > 0$ such that for all $\omega \epsilon S$, $\|x_0\| < \delta_\omega$ implies $\|g(t, x_0, t_0; \omega)\| < \epsilon$ for all $t \geq t_0$; that is, for almost every ω, the trivial solution is stable.

The dependence of δ_ω on ω is somewhat unsatisfactory for given t_0 and a bound ϵ on the solutions, there is no guarantee that one can find a common bound $\delta(\epsilon, t_0)$ on the initial values x_0 to insure that the norm of almost all realizations (i. e., for all ω belonging to some set S with $P(S) = 1$) remains within ϵ. Clearly $\underline{\delta} = \inf_{\omega \epsilon S} \delta_\omega$ is the largest possible δ;

if $\underline{\delta} = 0$, however, then no such δ exists. In general, for the systems considered in the following sections, $\underline{\delta} = 0$ and this property of uniformity over a set of probability 1 will not be discussed further. If, however, it suffices to have a set of probability arbitrarily close to, but not necessarily equal to, one, on which there is a uniform bound on the initial conditions, then it can be shown that this is equivalent to almost sure stability. This motivates the following:

DEFINITION (i-f): The trivial solution is called stable almost uniformly-in-ω if given any $\eta > 0$, t_0, there exists an ω-set $B(\eta, t_0)$ with $P(B) > 1 - \eta$ and given any $\epsilon > 0$, t_0, there corresponds a number $\delta(\epsilon, \eta, t_0) > 0$ such that $\|x_0\| < \delta$ implies $\|g(t, x_0, t_0; \omega)\| < \epsilon$ for all $t \geq t_0$ and all $\omega \epsilon B$.

THEOREM 1: The trivial solution of (2. 2) is stable almost surely if and only if it is stable almost uniformly-in-ω. (Proof: See Appendix C.)

It is clear that (i-f) implies (i-c); hence by the preceding theorem, (i-d) also implies (i-c).

The concept of uniform stability for deterministic systems has its natural counterpart in stochastic systems.

DEFINITIONS (ii-a)-(ii-f): The trivial solution is said to be uniformly stable in the mean norm, in the mean square, in the ith moment, in probability, almost surely, almost uniformly-in-ω, if it satisfies definitions (i-a), (i-b), (i-c), (i-d), (i-e), (i-f) respectively, with δ independent of t_0 (and B independent of t_0 in definition (i-f)).

The following theorem parallels Theorem 1:

THEOREM 2: The trivial solution of (2.2) is uniformly stable almost surely if and only if it is uniformly stable almost uniformly-in-ω.

PROOF: The proof follows at once from the proof of Theorem 1 upon noting the uniformity in t_0 of B and of δ.

ASYMPTOTIC CONSIDERATIONS. In the preceding section, the definitions of stability relate to the behavior of the solution function for all time $t \geq t_0$. In this section it is the large time or asymptotic behavior of the solution functions that is of interest. Additional requirements of appropriate behavior for all time $t \geq t_0$ may or may not be imposed.

The collection of solution functions $g(t, x_0, t_0; \omega)$ is a class, indexed by (x_0, t_0), of random processes. The convergence in t of the random processes for all (x_0, t_0) belonging to some set is of interest. The types of convergence considered are those commonly studied in probability theory, namely, in the ith mean, in probability, almost sure, and almost uniform (-in-ω). (See Loeve.[21]) From the usual definitions of these modes of convergence the definitions of stability follow, attention being paid to the dependence of the random processes on the couple (x_0, t_0).

DEFINITIONS: The trivial solution is called

(iii-a) quasi-asymptotically stable in the mean norm if given t_0 there exists $\delta(t_0) > 0$ such that $\| x_0 \| < \delta$ implies $E\| g(t, x_0, t_0)\| \longrightarrow 0$ as $t \longrightarrow \infty$.

(iii-b) quasi-asymptotically stable in the mean square if given t_0 there exists $\delta(t_0) > 0$ such that $\| x_0 \| < \delta$ implies

$$E \sum_{j=1}^{n} \left(g_j(t, x_0, t_0)\right)^2 \longrightarrow 0 \text{ as } t \to \infty.$$

(iii-c) quasi-asymptotically stable in the ith moment if given t_0 there exists $\delta(t_0) > 0$ such that $\|x_0\| < \delta$ implies $Eg_{[i]}(t, x_0, t_0) \to 0$ as $t \to \infty$.

(iii-d) quasi-asymptotically stable in probability if given $\eta > 0 > t_0$, there exists $\delta(\eta, t_0) > 0$ such that given any $\epsilon > 0$ there is a $T(\epsilon, \eta, x_0, t_0) > 0$ such that $\|x_0\| < \delta$ implies $P[\|g(t, x_0, t_0)\| < \epsilon] > 1 - \eta$ for all $t \geq t_0 + T$.

(iii-e) quasi-asymptotically stable almost surely if there exists an ω-set S with $P(S) = 1$ such that for every $\omega \epsilon S$ the trivial solution is quasi-asymptotically stable; or equivalently, there exists an ω-set S with $P(S) = 1$ such that given t_0, $\omega \epsilon S$, there exists $\delta(\omega, t_0) > 0$, and for any $\epsilon > 0$, there is a $T(\epsilon, \omega, x_0, t_0) > 0$ such that $\|x_0\| < \delta$ implies $\|g(t, x_0, t_0; \omega)\| < \epsilon$ for all $t \geq t_0 + T$.

(iii-f) quasi-asymptotically stable almost uniformly-in-ω if given $\eta > 0, t_0$, there exists $\delta(\eta, t_0) > 0$, and given any x_0, $\|x_0\| < \delta$, there is an ω-set $B(\eta, x_0, t_0)$ with $P(B) > 1-\eta$ such that $\|x_0\| < \delta$ implies $g(t, x_0, t_0; \omega) \to 0$ as $t \to \infty$ uniformly on B; or equivalently, given $\eta > 0$, t_0, there exists $\delta(\eta, t_0) > 0$ such that $\|x_0\| < \delta$ implies there exists an ω-set $B(\eta, x_0, t_0)$ with $P(B) > 1-\eta$ having the property that: to every $\epsilon > 0$, there corresponds a $T(\epsilon, \eta, x_0, t_0) > 0$ such that $\|x_0\| < \delta$ implies $\|g(t, x_0, t_0; \omega)\| < \epsilon$ for all $t \geq t_0 + T$ and all $\omega \epsilon B$.

The discussion given at the end of definition (ii-e) is equally pertinent here. In definition (iii-e) the bound on the initial value and the number T governing the rate of convergence depend on ω; in definition (iii-f) both δ and T are uniform in ω but this uniformity is over an ω-set B whose probability can be chosen to be arbitrarily close to, but not always equal to, one. The following two theorems complement Theorem 1.

THEOREM 3: The trivial solution of (2.2) is quasi-asymptotically stable almost surely if it is quasi-asymptotically stable almost uniformly-in-ω with the set B(see definition (iii-f) above) independent of x_0. (Proof: See Appendix C.)

THEOREM 4: The trivial solution of (2.2) is quasi-asymptotically stable almost uniformly-in-ω if it is quasi-asymptotically stable almost surely. (Proof: See Appendix C.)

Clearly (iii-f) implies (iii-c), whence by Theorem 4 (iii-d) also implies (iii-c).

The statements of the above theorems can be suitably altered to incorporate some requirement of uniformity of δ, B, T, in the variables x_0, t_0. The proofs would parallel closely those of Theorems 1-4. Definitions (i-d)-(i-f), (ii-d)-(ii-f), (iii-d)-(iii-f) can also be modified somewhat to describe closely related concepts. For example, definitions (i-e), (iii-e) may be altered so that the set S of unit probability depends on the choice of t_0. Thus the definitions and theorems given in this section are merely indicative of the concepts and results that can be obtained. They are by no means exhaustive.

In the deterministic case, asymptotic stability requires both stability and quasi-asymptotic stability. Likewise, in the random case, the following definitions can be formulated:

DEFINITION: The trivial solution is called

(iv-a) asymptotically stable in the mean norm if it is (ii-a) and (iii-a).

(iv-b) asymptotically stable in the mean square if it is (ii-b) and (iii-b).

(iv-c) asymptotically stable in the ith moment if it is (ii-c) and (iii-c).

(iv-d) asymptotically stable in probability if it is (ii-d) and (iii-d).

(iv-e) asymptotically stable almost surely if it is (ii-e) and (iii-e).

(iv-f) asymptotically stable almost uniformly-in-ω if it is (ii-f) and (iii-f).

It follows from their respective definitions that (iv-f) implies (iv-d); hence by the above theorem, (iv-e) also implies (iv-b).

The remaining concepts of stability, (v) - (ix), can be similarly generalized to the random case. To avoid repetitiousness, they will not be discussed here. Exponential asymptotic stability is, however, considered briefly for it is used specifically in Chap. V.

DEFINITION: The trivial solution is called

(ix-a) <u>exponentially asymptotically stable in the mean norm</u> if there exists a $\nu > 0$, and given any $\epsilon > 0$ there is a $\delta(\epsilon) > 0$ such that for all t_0, $\| x_0 \| < \epsilon$ implies $E \| g(t, x_0, t_0) \| < \epsilon \exp[-\nu(t-t_0)]$ for all $t \geq t_0$.

(ix-c) exponentially asymptotically stable in the ith moment if there exists a $\nu > 0$, and given any $\epsilon > 0$ there is a $\delta(\epsilon) > 0$ such that for all t_0, $\| x_0 \| < \delta$ implies $\| Eg_{[i]}(t, x_0, t_0) \| < \epsilon \exp] - \nu(t-t_0)]$ for all $t > t_0$.

As before, (ix-c) for even i implies (ix-a). If the system is linear, that is, if f in Eq. (2. 2) is almost surely linear in x, then (ix-a) implies uniform asymptotic stability in the mean norm and (ix-c) implies uniform asymptotic stability in the ith moment. Moreover, by virtue of the linearity in x of the solution functions of the differential system, for both (ix-a) and ix-c) there exists a constant $K \geq 1$ such that it is possible to choose $\delta(\epsilon) = \epsilon / K$ so that $E \| g(t, x_0, t_0) \| < K \| x_0 \| \exp[-\nu(t-t_0)]$ and $\| Eg_{[i]}(t, x_0, t_0) \| < K \| x_0 \| \exp[-\nu(t-t_0)]$ respectively.

III. LINEAR SYSTEMS WITH INDEPENDENT PARAMETERS

SUMMARY

This section is devoted to the stability analysis of the random linear differential system with piecewise constant coefficients:

$$\dot{x} = A_k x, \quad t_{k-1} \leq t < t_k, \quad k = 1 \ 2 \ \ldots \tag{3.1}$$

where x(t) is an n-vector and A_k is a constant n x n matrix. The behavior of the system is completely determined by the random sequence $\{A_k(t_k - t_{k-1})\}$ here assumed to be an independent, identically distributed, random process. Chapter IV treats the case where $\{A_k(t_k - t_{k-1})\}$

is a homogeneous Markov chain. * Necessary and sufficient conditions for stability, and (exponential) asymptotic stability, in the ith moment, are obtained. It is shown that almost sure asymptotic stability is a consequence of asymptotic stability in the ith moment, i an even integer, and that the converse implication is false.

3.1 ith Moment Stability

Given an initial vector $x(t_0) = x_0$ the solution of the differential equation (3.1) in the first time interval is

$$x(t) = e^{A_1(t-t_0)} x_0 , \qquad t_0 \leq t < t_1$$

Defining the solution function at time t_1 by continuity from the left,

$$x(t_1) = e^{A_1(t_1-t_0)} x_0 ,$$

and taking this to be the initial value for the next time interval, the solution in the second time interval is

$$x(t) = e^{A_2(t-t_1)} x(t_1)$$

$$= e^{A_2(t-t_1)} e^{A_1(t_1-t_0)} x_0 , \qquad t_1 \leq t < t_2$$

It follows that the time function

$$x(t) = e^{A_k(t-t_{k-1})} e^{A_{k-1}(t_{k-1}-t_{k-2})} \cdots e^{A_1(t_1-t_0)} x_0 , \tag{3.2}$$

$$t_{k-1} \leq t < t_k , \quad k = 1, 2, 3, \ldots$$

satisfies the initial condition $x(t_0) = x_0$, and satisfies the differential equation (3.1) everywhere except at the time instants $t_1, t_2, t_3, \ldots$. The function $x(t)$ defined by (3.2) will be taken as the unique solution of (3.1) satisfying the initial condition $x(t_0) = x_0$.

Define

$$\Phi_k = e^{A_k(t_k-t_{k-1})} \tag{3.3}$$

---

* All of the results of this chapter can be deduced from those of Chap.IV (and Appendix E) which treats the Markov case. It is preferable, however, from an expository point of view, to treat first the simpler case of independence.

Then the solution (3.2) can be written as

$$x(t) = e^{A_k(t-t_{k-1})} \Phi_{k-1} \Phi_{k-2} \cdots \Phi_1 x_0 \tag{3.4}$$

$$t_{k-1} \leq t < t_k, \ k = 1, 2, \ldots$$

Evaluating the solution x(t) at times t_k only, (3.4) yields

$$x_k = \Phi_k \Phi_{k-1} \cdots \Phi_1 x_0 \tag{3.5}$$

where

$$x(t_k) = x_k = \Phi_k x_{k-1} \tag{3.6}$$

It is clear from the definition of Φ_k, Eq. (3.3), that the behavior of x_k depends only on the probabilistic structure of the random sequence $\{A_k(t_k - t_{k-1})\}$. It is also clear from the continuity, with respect to the initial value x_0, of solutions of the differential equation (3.1) that it suffices to examine the sequence $\{x_k\}$ to determine the stability (of the trivial solution) of the differential system (3.1), provided the Φ_k are bounded in some sense. Hence it will be assumed that the Φ_k are bounded (almost surely) uniformly in k, i.e., there exists a finite M such that for all k, $P[\|\Phi_k\| < M] = 1$.

In the remainder of this section it is assumed that $\{\Phi_k\}$ is a sequence of independent, identically distributed, random matrices of order n x n. Since the sequence $\{x_k\}$ is generated by subjecting a given vector x_0 to a succession of independent, identically distributed, random linear transformations: $x_k = \Phi_k \Phi_{k-1} \cdots \Phi_1 x_0$, the random sequence $\{x_k\}$ is a homogeneous, nonstationary, infinite Markov chain. (See Chap. IV, Secs. 4.1 and 4.2.)

Take the expectation of both sides of Eq. (3.5) to obtain

$$Ex_k = E(\Phi_k \Phi_{k-1} \cdots \Phi_1) x_0,$$

which by independence becomes

$$Ex_k = (E\Phi_k)(E\Phi_{k-1}) \cdots (E\Phi_1) x_0$$

Moreover, since the Φ_k have a common distribution,

$$Ex_k = (E\Phi)^k x_0 , \tag{3. 7}$$

where the subscript on Φ has been dropped because $E\Phi_j$ does not depend on j.

Equation (3. 7) shows that the behavior of Ex_k as a function of k depends only on $E\Phi$ (for fixed x_0), in particular, only on the location and the index of the eigenvalues of $E\Phi$. This point is elaborated on later when it is shown that the computation of the ith moment, $i = 1, 2, 3, \ldots,$ of x_k leads to equations identical in form to (3. 7).

To evaluate the second moment it is necessary to work with the squares of the components of x_k. For algebraic simplicity assume that the system is order (n=) 2. Let $x_k = (u_k, v_k)$. Then the relation $x_k = \Phi_k x_{k-1}$ can be written as

$$\begin{bmatrix} u_k \\ v_k \end{bmatrix} = \begin{bmatrix} \phi_{11}^{(k)} & \phi_{12}^{(k)} \\ \phi_{21}^{(k)} & \phi_{22}^{(k)} \end{bmatrix} \begin{bmatrix} u_{k-1} \\ v_{k-1} \end{bmatrix}$$

which may be expressed as

$$u_k = \phi_{11}^{(k)} u_{k-1} + \phi_{12}^{(k)} v_{k-1}$$

$$v_k = \phi_{21}^{(k)} u_{k-1} + \phi_{22}^{(k)} v_{k-1}$$

Form all second-order products of components of x_k to obtain

$$
\begin{aligned}
u_k^2 &= \phi_{11}^{(k)2} u_{k-1}^2 + \phi_{11}^{(k)} \Phi_{12}^{(k)} u_{k-1} v_{k-1} + \phi_{12}^{(k)} \phi_{11}^{(k)} v_{k-1} u_{k-1} + \phi_{12}^{(k)2} v_{k-1}^2 \\
u_k v_k &= \phi_{11}^{(k)} \phi_{21}^{(k)} u_{k-1}^2 + \phi_{11}^{(k)} \phi_{22}^{(k)} u_{k-1} v_{k-1} + \phi_{12}^{(k)} \phi_{21}^{(k)} v_{k-1} u_{k-1} + \phi_{12}^{(k)} \phi_{22}^{(k)} v_{k-1}^2 \\
v_k u_k &= \phi_{21}^{(k)} \phi_{11}^{(k)} u_{k-1}^2 + \phi_{21}^{(k)} \phi_{12}^{(k)} u_{k-1} v_{k-1} + \phi_{22}^{(k)} \phi_{11}^{(k)} v_{k-1} u_{k-1} + \phi_{22}^{(k)} \phi_{12}^{(k)} v_{k-1}^2 \\
v_k^2 &= \phi_{21}^{(k)2} u_{k-1}^2 + \phi_{21}^{(k)} \phi_{22}^{(k)} u_{k-1} v_{k-1} + \phi_{22}^{(k)} \phi_{21}^{(k)} v_{k-1} u_{k-1} + \phi_{22}^{(k)2} v_{k-1}^2
\end{aligned}
\tag{3.8}
$$

Define $x_{k[2]}$ by

$$
x_{k[2]} = \begin{bmatrix} u_k^2 \\ u_k v_k \\ v_k u_k \\ v_k^2 \end{bmatrix}
$$

Then (3. 7) shows that $x_{k[2]}$ is a <u>linear transform</u> of $x_{k-1[2]}$,

$$
x_{k[2]} = \Phi_{k[2]} x_{k-1[2]}, \tag{3.9}
$$

the matrix $\Phi_{k[2]}$ of the transformation being defined by the set of equations (3. 8). The recursion equation (3. 9) is linear, as is the original equation (3. 6), yet it relates second-order forms of components of x_k, this being achieved at the cost of working in a higher dimension. The technique used is, in fact, that of Kronecker products (as explained below), which dates back to the last century (see MacDuffee,[22] pp. 81-86) and which was used recently by Bellman[14] in solving a similar problem of computing moments.

The matrix $\Phi_{k[2]}$ in Eq. (3.9) is recognized to be the Kronecker (or direct) product of Φ_k by itself:

$$\Phi_{k[2]} = \Phi_k \otimes \Phi_k;$$

likewise, $x_{k[2]}$ is the Kronecker product of x_k by itself:

$$x_{k[2]} = x_k \otimes x_k$$

(Appendix A is a brief outline of some of the useful properties of Kronecker and power products of matrices.)

By using the theory of Kronecker products, Eq. (3.6), valid for systems of <u>any</u> order, can be easily derived. Take the self-Kronecker product on both sides of Eq. (3.6) to obtain

$$x_k \otimes x_k = (\Phi_k x_k) \otimes \Phi_k x_k \tag{3.10}$$

where x_k is an n-vector. Now, for any two matrices A and B, not necessarily square but such that the product AB is defined,

$$AB \otimes AB = (AB)_{[2]} = A_{[2]} B_{[2]} = (A \otimes A)(B \otimes B); \tag{3.11}$$

that is, the operations of the Kronecker product and the conventional matrix product commute. It is this commutative property that makes the theory of Kronecker products so applicable to the problem of evaluating moments. For it follows from Eq. (3.11) that (3.10) can be reduced to

$$x_{k[2]} = \Phi_{k[2]} x_{k-1[2]} \quad ,$$

whence by iteration on k

$$x_{k[2]} = \Phi_{k[2]} \Phi_{k-1[2]} \cdots \Phi_{1[2]} x_{0[2]}$$

Similarly, for ith-order products, form i-fold products from Eq. (3.6),

$$x_k \otimes x_k \otimes \cdots \otimes x_k = \Phi_k x_{k-1} \otimes \Phi_k x_{k-1} \otimes \cdots \otimes \Phi_k x_{k-1},$$

and by repeatedly using the commutative property (3.11), obtain

$$x_{k[i]} = \Phi_{k[i]} \Phi_{k-1[i]} \cdots \Phi_{1[i]} x_{0[i]} , \tag{3.12}$$

where the subscript [i] denotes the ith self-Kronecker product.

Now, $\{\Phi_{k[i]}\}$ is an independent, identically distributed, random sequence of matrices (of order n^i). Hence, taking the expectation in Eq. (3.12),

$$Ex_{k[i]} = \left(E\Phi_{[i]}\right)^k x_{0[i]} \tag{3.13}$$

The vector $x_{k[i]}$ has as its components all ith-order products of components x_k; hence, Eq. (3.13) gives an explicit and simple relationship for studying the ith moment of components of x_k. The knowledge of the moments at time t_k may be used to study the probability distribution of x_k (see Cramer,[22] p. 176). Also, bounds on the probability of deviation from zero of any component of x_k may be evaluated readily by using the Markov inequality[21]

$$P[\,|Z| \geq \epsilon\,] \leq \frac{E|Z|^r}{\epsilon^r}, \qquad \epsilon > 0, \tag{3.14}$$

where Z is a scalar random variable and r is a positive real number.

Equation (3.13) shows that the behavior of large powers of a matrix is of interest. Hence consider the following

LEMMA: Let F be a complex matrix of finite order. Then

(i) $\lim_{k\to\infty} F^k = 0$ if and only if every eigenvalue of F lies inside the unit circle; in fact, $\|F^k\|$ will converge exponentially: there exist two constants $K_1 \geq 1$, $\nu_1 > 0$ such that $\|F^k\| < K_1 \exp[-\nu_1 k]$;

(ii) $\lim_{k \to \infty} F^k = F_1 \neq 0$ if and only if one is an eigenvalue of F of index* one and all other eigenvalues of F are inside the unit circle;

(iii) F^k is bounded uniformly in k but does not converge as $k \to \infty$ (is "oscillatory") if and only if F has no eigenvalues outside the unit circle, all eigenvalues on the unit circle are of index one, and there exists an eigenvalue not equal to one on the unit circle;

(iv) $\lim_{k \to \infty} \| F^k \| = \infty$ if and only if F has an eigenvalue outside the unit circle and/or has eigenvalues of index greater than one on the unit circle; in fact, $\| F^k \|$ will diverge exponentially: there exist two constants $K_2 \geq 1$, $\nu_2 > 0$ such that $\| F^k \| > K_2 \exp[+\nu_2 k]$.

PROOF: To arrive at the above conclusions, it is only necessary to transform F into a Jordan canonical form so that powers of F can be studied readily.

THEOREM 1: The trivial solution is stable in the ith moment (A) if and (A') only if no eigenvalue of $E\Phi_{[i]}$ lies outside the unit circle and any eigenvalue on the unit circle has index one. It is asymptotically stable in the ith moment (B) if and (B') only if all eigenvalues of $E\Phi_{[i]}$ lie inside the unit circle. Moreover, in (A) and (B) it is uniformly stable in the ith moment and uniformly asymptotically stable in the ith moment respectively.

PROOF: (A) By the assumptions on the eigenvalues and by the above Lemma, $\sup_{k \geq 1} \| (E\Phi_{[i]})^k \| < M$ for some finite M. Given any $\epsilon > 0$, choose $\delta_i(\epsilon) = \epsilon / M$. Then by Eq. (3.13),

$$\| Ex_{k[i]} \| \leq \| (E\Phi_{k[i]})^k \| \; \| x_{0[i]} \| < M \| x_{0[i]} \| . \tag{3.15}$$

If $\| x_{0[i]} \| < \delta_i$, then $\| Ex_{k[i]} \|$ is less than ϵ. But $\| x_{0[i]} \| < \delta_i$ is equivalent to $\| x_0 \| < \delta$ for some $\delta > 0$. Furthermore, δ', and hence δ, do not depend on t_0. Hence the trivial solution is uniformly stable in the ith moment.

(A') If the eigenvalue assumption is violated, then by part (iv) of the above Lemma, $\|(E \Phi_{[i]})^k\| \to \infty$ as $k \to \infty$; or equivalently, some element of the matrix $(E\Phi_{[i]})^k$ becomes unbounded as $k \to \infty$.

---

*The index of an eigenvalue is its multiplicity as a root of the minimal polynomial of F. See also Friedman,[24] Chap. 2.

Hence there does not exist any $\delta > 0$ with the property that for some t_0, $\| Ex_{k[i]} \| = \|(E\Phi_{[i]})^k x_{0[i]}\|$ remains bounded as $k \to \infty$ for <u>all</u> $\| x_0 \| < \delta$. This implies that the trivial solution cannot be stable in the ith moment.

(B) By part (i) of the preceding Lemma, $(E\Phi_{[i]})^k \to 0$ as $k \to \infty$. Clearly this convergence is uniform in t_0. Hence, given any $\delta > 0$, it follows by Eq. (3.15) that for all t_0, $\| x_0 \| < \delta$ implies $Ex_{k[i]} \to 0$ as $k \to \infty$ uniformly in t_0, and uniformly on $\| x_0 \| < \delta$ (since the system is linear). Hence the trivial solution is uniformly quasi-asymptotically stable in the ith moment. Moreover by (A), it is uniformly stable in the ith moment. Therefore it is uniformly asymptotically stable in the ith moment.

(B') If every eigenvalue of $E\Phi_{[i]}$ is not inside the unit circle, then by part (i) of the preceding Lemma, $(E\Phi_{[i]})^k$ will not converge to the zero matrix as $k \to \infty$. Hence there exists no $\delta > 0$ such that for some t_0, $Ex_{k[i]} = (E\Phi_{[i]})^k x_{0[i]} \to 0$ as $k \to \infty$ for <u>all</u> $\| x_0 \| < \delta$. Therefore the trivial solution cannot be quasi-asymptotically stable in the ith moment and hence asymptotically stable in the ith moment.

For subsequent use, especially in Chap. V in connection with the inhomogeneous (forced) random linear system, it is useful to establish the property of exponential asymptotic stability:

THEOREM 2: For the random linear system (3.1), the following propositions are equivalent:

(a) All eigenvalues of $E\Phi_{[i]}$ are inside the unit circle.

(b) The trivial solution is uniformly asymptotically stable in the ith moment.

(c) The trivial solution is exponentially asymptotically stable in the ith moment.

PROOF: The equivalence of (a) and (b) was established in Theorem 1.

That (c) implies (b) follows at once from the definitions of exponential and uniform asymptotic stability in the ith moment.

To show that (a) implies (c): Since all eigenvalues of $E\Phi_{[i]}$ are inside the unit circle, by part (iv) of the Lemma there are two constants $K \geq 1$ and $\nu > 0$ such that $\| (E\Phi_{[i]})^k \| < K \exp(-\nu k)$. Hence by Eq. (3.13),

$$\| Ex_{k[i]} \| < K \| x_{0[i]} \| \exp(-\nu k) < K_i \| x_0 \| \exp(-\nu k)$$

for some $K_i \geq 1$ from which follows the desired result.

An inspection of the set of equations (3.8) reveals that the third equation is superfluous and leads to the condensed set of equations

$$\begin{bmatrix} u_k^2 \\ u_k v_k \\ v_k^2 \end{bmatrix} = \begin{bmatrix} \phi_{11}^{(k)2} & 2\phi_{11}^{(k)}\phi_{12}^{(k)} & \phi_{12}^{(k)2} \\ \phi_{11}^{(k)}\phi_{21}^{(k)} & \phi_{11}^{(k)}\phi_{22}^{(k)}+\phi_{12}^{(k)}\phi_{21}^{(k)} & \phi_{12}^{(k)}\phi_{22}^{(k)} \\ \phi_{21}^{(k)2} & 2\phi_{21}^{(k)}\phi_{22}^{(k)} & \phi_{22}^{(k)2} \end{bmatrix} \begin{bmatrix} u_{k-1}^2 \\ u_{k-1}v_{k-1} \\ v_{k-1}^2 \end{bmatrix}$$

which may be expressed as

$$x_{k(2)} = \Phi_{k(2)} x_{k-1(2)}$$

The vector $x_{k(2)}$ and the matrix $\Phi_{k(2)}$ are recognized to be the self-power products (see Appendix A) of x_k and Φ_k, respectively. As in the Kronecker product case, for a system of any order,

$$Ex_{k(i)} = (E\Phi_{(i)})^k x_{0(i)}$$

is the relation for the ith moment and the eigenvalues of $E\Phi_{(i)}$ determine the asymptotic behavior of the ith moment.

In theoretical analysis, the symmetrical structure of the Kronecker product matrix is most useful; on the other hand, the smaller order of the corresponding power product matrix can be distinctly advantageous in reducing the numerical work. The Kronecker product will be used exclusively hereafter.

A central problem in the application of the above methods is the study of the location of the eigenvalues of matrices with respect to the unit circle. Either Schur-Cohn or transformed Routh-Hurwitz criteria may be used (see Marden,[25] Bharucha,[26] Jury[26-28]) but paper-and-pencil methods are tedious even for systems of low order. The problem, however, readily lends itself to solution on a digital computer.

3.2 Almost Sure Asymptotic Stability

By Eq. (3.13) and Theorem 2, if all the eigenvalues of $E\Phi_{[i]}$ are inside the unit circle, then the trivial solution is asymptotically stable in the ith moment and $Ex_{k[i]}$ converges exponentially to zero. The exponential convergence implies that $\sum_{k=1}^{\infty} Ex_{k[i]}$ exists and is finite, which via the Markov inequality and the Borel-Cantelli lemma shows that $x_k \xrightarrow{a.s.} 0$ if i is even. Since the system is linear, this convergence is uniform in x_0, from which follows

THEOREM 3: The trivial solution is equiasymptotically stable almost surely if for some even i, all eigenvalues of $E\Phi_{[i]}$ are inside the unit circle.

PROOF: By Eq. (3.13)

$$Ex_{k[i]} = (E\Phi_{[i]})^k x_{0[i]}$$

Summing on k,

$$\sum_{k=1}^{\infty} Ex_{k[i]} = \sum_{k=1}^{\infty} (E\Phi_{[i]})^k x_{0[i]} = (I - E\Phi_{[i]})^{-1} x_{0[i]} \qquad (3.16)$$

since all eigenvalues of $E\Phi_{[i]}$ have moduli less than one.

Let u_k be any component of x_k. Then Eu_k^i is a component of $Ex_{k[i]}$, and by (3.15), $\sum_{k=1}^{\infty} Eu_k^i < \infty$. But by the Markov inequality, (3.14),

$$P[\,|u_k| \geq \epsilon_1] \leq \frac{E|u_k|^i}{\epsilon_1^i}, \quad \epsilon_1 \geq 0.$$

Hence $\sum_k P[\,|u_k| \geq \epsilon_1] < \infty$ for every $\epsilon_1 > 0$, which by the Borel-Cantelli lemma implies that $u_k \xrightarrow{a.s.} 0$.

Therefore, given any $\delta > 0$, for all t_0, $\|x_0\| < \delta$ implies $g(t, x_0, t_0) \longrightarrow 0$. Moreover, since the system is linear the convergence to zero is uniform on $\|x_0\| < \delta$. Hence the trivial solution is quasi-equiasymptotically stable almost surely. Also, for almost every ω, given any $\epsilon > 0$, t_0, there exists a $T(\epsilon, \delta, t_0;\omega) > 0$ such that $\|x_0\| < \delta$ implies $\|g(t, x_0, t_0; \omega)\| < \epsilon$ for all $t \geq t_0 + T$. Now T is independent of x_0, $\|x_0\| < \delta$, and $g(t, x_0, t_0;\omega)$ is continuous in x_0; hence $\max_{t_0 \leq t \leq t_0 + T} \|g(t, x_0, t_0; \omega)\|$ is continuous in x_0. Therefore there exists a $\delta'(\epsilon, t_0) > 0$, $\delta' < \delta$, such that $\|x_0\| < \delta'$ implies $\|g(t, x_0, t_0;\omega)\| < \epsilon$ for all $t \geq t_0 + T$ (and also for $t_0 \leq t \leq t_0 + T$ by definition of δ' and T). Consequently the trivial solution is stable almost surely. (This also could have been derived by using the linearity in x_0, instead of using the continuity in x_0, of the solution function.) Since the trivial solution is, in addition, quasi-equiasymptotically stable almost surely, it is equiasymptotically stable almost surely. Note that the only reason the solution is not <u>uniformly</u> asymptotically stable almost surely is that T depends on t_0, that is, the convergence to zero is not necessarily uniform in t_0.

The converse of Theorem 3 is not true: Example 1 below shows that the trivial solution can be almost surely quasi-asymptotically stable although for all i, $E\Phi_{[i]}$ has eigenvalues outside the unit circle. This is in contrast to Theorem 1 which states that asymptotic stability in the ith moment is equivalent to having all eigenvalues of $E\Phi_{[i]}$ inside the unit circle.

3.3 Example 1

Let $t_k - t_{k-1} = T$ (fixed)

$A_k = b_1 \leq 0$ with probability p

$A_k = b_2 \geq 0$ with probability 1-p .

Then

$$\Phi_k = e^{b_1 T} \quad \text{with probability } p$$

$$\Phi_k = e^{b_2 T} \quad \text{with probability } 1-p$$

and

$$\Phi_{k[i]} = e^{ib_1 T} \quad \text{with probability } p$$

$$\Phi_{k[i]} = e^{ib_2 T} \quad \text{with probability } 1-p.$$

Hence $E\Phi_{[i]} = pe^{ib_1 T} + (1-p)e^{ib_2 T}$. By Theorem 1, the trivial solution is stable in the ith moment if and only if

$$p \geq \frac{e^{ib_2 T} - 1}{e^{ib_2 T} - e^{ib_1 T}}$$

It is asymptotically stable in the ith moment if the above constraint is changed to a strict inequality. These results are exhibited graphically in Fig. 1. Note that the region of stability increases as b_1, b_2 decrease. This is to be expected for b_1^{-1}, b_2^{-1} are the respective "time constants" of the stable and unstable modes of the system.

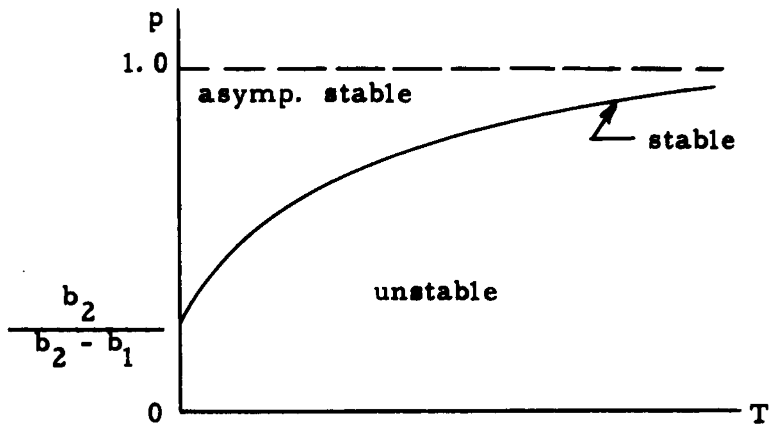


Figure 1

It is interesting to examine the situation when $b_1 = -\infty$, $\Phi_k = e^{b_1 T} = 0$, $T \neq 0$, with probability p. The curve in Fig. 1 delineating the region of stability now starts at the origin and rises exponentially to one. But for all $p > 0$, $T > 0$, the trivial solution is a. s. asymptotically stable because the only realization that does not become zero as $t \to \infty$ is the one where the "stable" mode $e^{b_1 T}$ is never achieved and thus has probability $(1-p)^{\infty} = 0$. However, it should be noted that for any $p < 1$, there exists a positive probability that the trivial solution is unbounded; for, given any bound $M > 0$, $\|x(t)\|$ will exceed M if the system starts in and remains in the unstable mode for a finite length of time.

3.4 Example 2

The time intervals $t_k - t_{k-1}$ are of fixed length T. The system has a double "pole" at $a \leq 0$ and $b \geq 0$ with probabilities p and 1-p, respectively:

$$A_k = \begin{bmatrix} a & T \\ 0 & a \end{bmatrix} \quad \text{with probability } p$$

$$A_k = \begin{bmatrix} b & T \\ 0 & b \end{bmatrix} \quad \text{with probability } 1-p$$

The corresponding values of Φ_k are $e^{aT}B$ and $e^{bT}B$ where

$$B = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$$

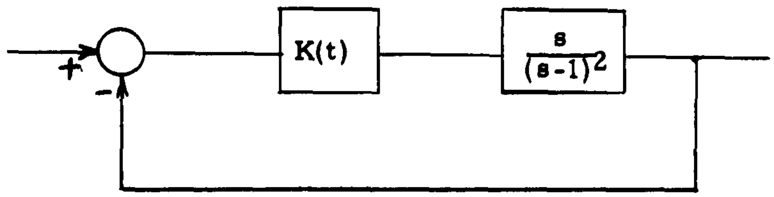


Figure 2

For example, the feedback control system shown above would have $a = -1$, $b = 1$ if the random gain K was equal to 4 or 0 with respective probabilities p and 1-p.

Returning to the general case,

$$E\Phi = (pe^{aT} + (1-p)e^{bT})\, B$$

B has a double eigenvalue at +1. Hence the eigenvalues of $E\Phi$ are inside the unit circle if and only if

$$pe^{aT} + (1-p)e^{bT} < 1,$$

that is, if and only if

$$p > \frac{e^{bT} - 1}{e^{bT} - e^{aT}}$$

in which case $Ex_{\infty} = 0$.

Again,

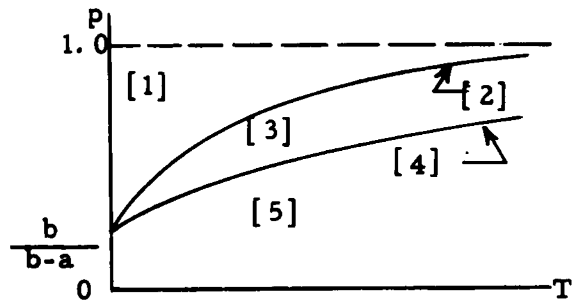
$$E\,\Phi_{[2]} = (pe^{2aT} + (1-p)e^{2bT})\, B_{[2]}$$

Note that +1 is an eigenvalue of B of multiplicity two. Hence, +1 is an eigenvalue of $B_{[2]}$ of multiplicity four (see Appendix A, Eq. A. 5). Moreover, it can be shown that its index is also four.

The eigenvalues of $E\Phi_{[2]}$ are inside the unit circle if and only if

$$p > \frac{e^{2bT} - 1}{e^{2bT} - e^{2aT}} \tag{3.17}$$

in which case $Ex_{k[2]} \longrightarrow 0$ as $k \longrightarrow \infty$ and the system is asymptotically stable in the second moment. If (3.17) is changed to an equality, the system becomes unstable in the second moment for $E\Phi_{[2]}$ now has one as an eigenvalue of index four on the unit circle. Figure 3 summarizes the above results.



Asymptotically stable in the second moment: [1]
Unstable in the second moment: [2] - [5]
Ultimate mean zero: [1] - [3]
Ultimate mean infinite: [4] - [5]

Figure 3

Suppose a = -1 and b = 0. Then this example reduces to Example 1 of Bertram and Sarachik[3] where by Lyapunov's second method, the authors arrive at the result that $p > 1/(1-e^{-T})$ is a sufficient condition for asymptotic stability in the mean. But Eq. (3.17) implies that if $a < 0$, $b = 0$, a necessary and sufficient condition for asymptotic stability in the mean is $p > 0$. Indeed, by using the results of the next section, it can be shown that the latter result is true even if the sequence $\{\Phi_k\}$ is any finite-order Markov chain with a nonzero transition probability of going from the unstable to the stable mode.

IV. LINEAR SYSTEM WITH MARKOV PARAMETERS

Summary

The problem considered here is identical to the one treated in Chap. III except that the succession of modes of operation of the random linear system forms a Markov chain rather than just an independent, identically distributed, random sequence. The differential equation, as before, is

$$\dot{x} = A_k x, \qquad t_{k-1} \leq t < t_k, \qquad k = 1, 2, \ldots$$

where x(t) is an n-vector and A_k is a constant n x n matrix. The system behavior is completely described by the random sequence $\{A_k(t_k - t_{k-1})\}$, here assumed to be a Markov chain. Sufficient conditions for stability, and (exponential) asymptotic stability, in the ith moment, are obtained. It is shown that these conditions without further qualifications are not also necessary. The dependence of the ith moment of the solution function on the initial distribution and on the types of modes (states) of the Markov chain is discussed. It is shown that the sufficient conditions for asymptotic stability in the ith moment, i an even integer, also imply almost sure asymptotic stability.

4.1 Markov Chains

This section outlines some of the pertinent facts of the theory of Markov chains. Unfortunately there exists no standard terminology in the literature of Markov chains; a major purpose of this section is to introduce the notation and terminology to be used later in the section. Some general references on Markov chains are Bharucha-Reid,[29] Doob,[30] Gantmacher,[31] and Kemeny and Snell.[32]

A Markov process is an indexed set $\{Z_t, \ t \in \mathcal{T} \subset (-\infty, \infty)\}$ of random variables such that for any integer $n > 1$, any set $\{t_1 < t_2 < \ldots < t_n\}$ of parameter values, and any real number a,

$$P[Z_{t_n} \leq a \mid Z_{t_{n-1}}, \ldots, Z_{t_2}, Z_{t_1}] = P[Z_{t_n} \leq a \mid Z_{t_{n-1}}] \tag{4.1}$$

If

$$P[Z_{t_n} \leq a \mid Z_{t_{n-1}}, \ldots, Z_{t_1}] = P[Z_{t_n} \leq a \mid Z_{t_{n-1}}, \ldots, Z_{t_{n-\mu}}], \quad n > \mu, \tag{4.2}$$

then $\{Z_t\}$ is called a <u>Markov process of order</u> μ. By definition, if the term Markov process is used without any qualifications, the order is unity.

A random process is called a <u>discrete</u> or a <u>continuous parameter</u> process according to whether the index set $\mathcal{T}$ is countable or not. A random process is called <u>stationary</u> if for any $t_i \epsilon \mathcal{T}$, $i = 1, \ldots, n$, and any real α such that $(t_i + \alpha) \epsilon \mathcal{T}$, $i = 1, \ldots, n$, the joint distribution of the random variables $Z_{t_{1+\alpha}}$, $Z_{t_{2+\alpha}}, \ldots, Z_{t_{n+\alpha}}$ is independent of α for all finite n. The parameter t, sometimes, will be identified with the physical variable time.

If the random variables Z_t of a Markov process can assume values only in some countable set D, the process is called a <u>Markov chain</u>; the chain is said to be <u>finite</u> or <u>infinite</u> according as D is finite or infinite. If D is taken to be a collection of modes * D_j, the time development of a particular sample function of a discrete parameter Markov chain can be thought of as an evolution through the modes $D_j \epsilon D$:

$$\ldots, D_{j_1}, D_{j_2}, D_{j_3}, \ldots$$

For a discrete parameter Markov chain $\{Z_n, n = 1, 2, \ldots\}$, given an initial mode, the time development is completely described by a set of transition probabilities

$$p_{ij}(n) = P[Z_n = D_j \mid Z_{n-1} = D_i]$$

*The standard terminology of Markov chain theory calls for the use of the term "state" rather than "mode." The term "state" is not used here to prevent confusion with the state (vector) of the differential system.

specified for all $n > 1$, and all D_i, $D_j \in D$. Of course,

$$p_{ij}(n) \geq 0 \quad \text{and} \quad \sum_j p_{ij}(n) = 1$$

where the summation extends over all modes in D. The matrix

$$P(n) = (p_{ij}(n))$$

is called a Markov, or stochastic, or probability, or transition, matrix. Such a matrix together with a starting distribution

$$a_j = P[Z_1 = D_j]$$

completely specifies the chain. If p_{ij} is not a function of n, the chain is called homogeneous.

Henceforth, unless explicitly indicated, the term Markov chain will be used to denote finite homogeneous chains only.

If the elements of the kth power of the transition matrix of a Markov chain are denoted by $p_{ij}^{(k)}$,

$$P^k = (p_{ij}^{(k)}),$$

then

$$p_{ij}^{(k)} = P[Z_k = D_j \mid Z_{k-1} = D_i], \quad k > 1;$$

that is, $p_{ij}^{(k)}$ is the transition probability of going from mode D_i to mode D_j in k steps.

If for some finite $k \geq 1$, $p_{ij}^{(k)} > 0$, then it is possible to go (i. e., with nonzero probability) from mode i to mode j (the modes will often be referred to by their subscripts only), and this will be denoted by $i \rightarrow j$. If there exists no finite $k \geq 1$ for which $p_{ij}^{(k)} > 0$, that is, $p_{ij}^{(k)} = 0$ for all $k \geq 1$, then the transition from i to j cannot occur and this will be denoted by $i \not\rightarrow j$.

A transient set $\mathcal{T}$ of modes is the set of all modes i to which there correspond modes j such that $i \rightarrow j$ and $j \not\rightarrow i$. A nonempty set $\mathcal{K}$ of modes is called closed if for every $i \in \mathcal{K}$ and every $j \notin \mathcal{K}$, $i \not\rightarrow j$. If the jth row and jth column of some Markov matrix are deleted for all $j \notin \mathcal{K}$ where $\mathcal{K}$ is some closed set, the matrix remaining after deletion is still a Markov matrix. A mode i is called absorbing if and only if $p_{ii} = 1$; if a closed set contains only one mode then this must be an absorbing mode. A closed set $\mathcal{E}$ of modes is called ergodic if $i \rightarrow j$ for all i, $j \in \mathcal{E}$, i. e., $\mathcal{E}$ is a closed set no proper subset of which is closed. Every finite Markov chain contains at least one ergodic set; the complement of the union of all the ergodic sets is the transient set. If the set of all modes of a chain is ergodic, then the chain is called ergodic. A regular chain is an ergodic chain such that for all i, j, $p_{ij}^{(n)} > 0$ for all $n \geq N$ for some N. Every ergodic set can be partitioned into classes $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_d$, such that every one-step transition carries the chain from a mode in $\mathcal{C}_1$ to a mode in $\mathcal{C}_2$, from a mode in $\mathcal{C}_2$ to a mode in $\mathcal{C}_3, \ldots$, from a mode in $\mathcal{C}_d$ to a mode in $\mathcal{C}_1$; that is, the system moves cyclically through the classes ..., $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_d, \mathcal{C}_1, \ldots$. The integer d is called the period. If d = 1 then the ergodic set is called aperiodic; if d > 1 then the set is called cyclic. It follows that if $p_{ii}^{(n)} > 0$, $i \in \mathcal{E}$, then n is an integer multiple of d. It can be shown that there exists an integer b with the property that for all $i \in \mathcal{E}$, $p_{ii}^{(kd)} > 0$ for all $k \geq b$.

By a renumbering of the modes of a Markov chain, the transition matrix P always can be written in the form[31]

$$P = \begin{bmatrix} E_1 & 0 & \cdots & 0 & 0 \\ 0 & E_2 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & E_e & 0 \\ S_1 & S_2 & \cdots & S_e & T \end{bmatrix} \qquad (4.3)$$

$E_1, E_2, \ldots, E_e$ are themselves Markov matrices and correspond to the ergodic sets $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_e$ of the chain. The matrix T corresponds to the transient set. By a further reordering, any of the Markov matrices E_k can be written as[31]

$$E = \begin{bmatrix} 0 & C_1 & 0 & \cdots & 0 \\ 0 & 0 & C_2 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & C_{d-1} \\ C_d & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{4.4}$$

where p_{ij} is an element of the matrix C_n if $i \in \mathcal{C}_n$.

4.2 ith Moment Stability

As in Chap. III, the differential system is

$$\dot{x} = A_k x, \quad t_{k-1} \leq t < t_k, \qquad k = 1, 2, \ldots \tag{4.5}$$

which, as shown in Eqs. (3.1) to (3.6), may be reduced to

$$x_k = \Phi_k \Phi_{k-1} \cdots \Phi_1 x_0 \tag{4.6}$$

In this Chapter the matrix sequence $\{\Phi_k\}$ is a Markov chain with m possible modes $X_1, \ldots, X_m$, an initial probability distribution $\{a_i\}$ on the modes, and an m x m constant transition matrix $P = (p_{ij})$:

$$a_i = P[\Phi_1 = X_i] \qquad p_{ij} = P[\Phi_k = X_j \mid \Phi_{k-1} = X_i]$$

It is required to evaluate explicitly all the moments of the output vector process $\{x_k\}$, in particular, the limiting behavior as $k \to \infty$.

Now $x_k = \Phi_k x_{k-1} = \Phi_k \Phi_{k-1} x_{k-1}$. Hence, if x_{k-2} and x_{k-1} are observed, at say sample point ω, they will delineate a subset $\mathcal{R}_1$ of the range $\mathcal{R} = \{X_1, \ldots, X_m\}$ in which Φ_k must assume its value. $\mathcal{R}_1$, in general, is not coincident with $\mathcal{R}$. Since $\{\Phi_k\}$ is a first-order Markov chain, the distribution of Φ_k is conditioned on the subset $\mathcal{R}_1$. Hence $\{\Phi_k\}$ is, in general, a Markov process of order ≥ 1. Moreover, the range of $\{x_k\}$ is countable. So finally, $\{x_k\}$ is a homogeneous, nonstationary, infinite Markov chain of order ≥ 1.

Taking the expectation in Eq. (4.6),

$$Ex_k = E(\Phi_k \cdots \Phi_1)\, x_0$$

which by the derivation in Appendix D reduces to

$$Ex_k = \mathcal{J}_1 (Y_1 Q_1^T)^{k-1} Y_1 \mathcal{a}_1 x_0 \tag{4.7}$$

where Y_1 is the direct sum of the X_i:

$$Y_1 = X_1 \oplus \ldots \oplus X_m;$$

Q_1 is the square matrix of order mn:

$$Q_1 = P \otimes I_n, \quad I_n = \text{nth-order identity matrix;}$$

$\mathcal{a}_1$ is the mn x n matrix:

$$\mathcal{a}_1 = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \otimes I_n;$$

and $\mathcal{J}_1$ is the n x mn matrix:

$$\mathcal{J}_1 = [\underbrace{I_n, \ldots, I_n}_{\text{m identities}}]$$

Equation (4.7) gives the expected value of the sequence $\{x_k\}$. To obtain the ith moment, take the ith self-Kronecker product of Eq. (4.6) with itself to obtain (as derived in Eqs. (3.9) - (3.12))

$$x_{k[i]} = \Phi_{k[i]} \cdots \Phi_{1[i]} x_{0[i]}$$

Hence

$$Ex_{k[i]} = \mathcal{J}_i (Y_i Q_i)^{k-1} Y_i \mathcal{A}_i x_{0[i]} \tag{4.8}$$

where

$$Y_i = X_{1[i]} \oplus \cdots \oplus X_{m[i]} \ ,$$

$$Q_i = P \otimes I_{n^i} \quad ,$$

$$\mathcal{A}_i = \begin{bmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_m \end{bmatrix} \otimes I_{n^i} \quad ,$$

$$\mathcal{J}_i = [\underbrace{I_{n^i}, \ldots, I_{n^i}}_{\text{m identities}}]$$

If the Markov chain specializes to an independent process, Eq. (4.8) gives the results obtained in the previous chapter, as is shown in Appendix E. From Eq. (4.8), by using arguments similar to those in the previous chapter, sufficient conditions for stability and asymptotic stability, in the ith moment, can be found; that these conditions are not necessary is demonstrated later.

THEOREM 1: (A) The trivial solution of (4.5) is uniformly stable in the ith moment if $Y_i Q_i^T$ has no eigenvalues outside the unit circle, and any eigenvalue on the unit circle has index one. (B) It is uniformly asymptotically (in fact, exponentially asymptotically) stable in the ith moment if all eigenvalues of $Y_iQ_i^T$ lie inside the unit circle.

PROOF: (A) By the assumptions on the location and the index of the eigenvalues, it follows from the Lemma, Chap. III, that every element of the matrix $(Y_iQ_i^T)^k$ is bounded uniformly in k. Hence $\| \mathcal{J}_i(Y_iQ_i^T)^k Y_i \mathcal{a}_i \|$ is bounded uniformly in k, i.e., there exists an $M > 0$ such that $M = \sup_{k \geq 1} \| \mathcal{J}_i(Y_iQ_i^T)^k Y_i \mathcal{a}_i \|$. Given any $\epsilon > 0$, choose $\delta_i = \epsilon / M$. From Eq. (4.8)

$$\| Ex_{k[i]} \| \leq \| \mathcal{J}_i(Y_iQ_i)^{k-1} Y_i \mathcal{a}_i \| \| x_{0[i]} \| < M \| x_{0[i]} \| . \tag{4.9}$$

If $\| x_{0[i]} \| < \delta_i$, then $\| Ex_{k[i]} \|$ is less than ϵ. But $\| x_{0[i]} \| < \delta_i$ is equivalent to $\| x_0 \| < \delta$ for some $\delta > 0$. Also δ_i, and hence δ, do not depend on t_0 because the starting probability distribution is taken to be independent of t_0. Hence the trivial solution is uniformly stable in the ith moment.

(B) Since all eigenvalues of $Y_iQ_i^T$ have moduli less than one, it follows by the Lemma, Chap. III, that $(Y_iQ_i^T)^k \longrightarrow 0$ as $k \rightarrow \infty$. Hence $\mathcal{J}_i(Y_iQ_i^T)^k Y_i \mathcal{a}_i \rightarrow 0$ as $k \rightarrow \infty$. Note, moreover, that the convergence to the zero matrix is independent of t_0 for the Markov chain is assumed to be homogeneous with an initial distribution that is independent of t_0. Therefore, given any $\delta > 0$, it follows from Eq. (4.9) that for all t_0, $\| x_0 \| < \delta$ implies $Ex_{k[i]} \longrightarrow 0$ as $k \rightarrow \infty$ uniformly in t_0, and uniformly on $\| x_0 \| < \delta$ (since the system is linear). Hence the trivial solution is uniformly quasi-asymptotically stable in the ith moment. Moreover by (A) it is uniformly stable in the ith moment. Therefore it is uniformly asymptotically stable in the ith moment.

To show exponential asymptotic stability in the ith moment, take norms in Eq. (4.8) to obtain

$$\| Ex_{k[i]} \| \leq \| \mathcal{G}_i \| \; \| (Y_i Q_i^T)^{k-1} \| \; \| Y_i \mathcal{a}_i \| \; \| x_{0[i]} \|$$

and then use the fact that $\| (Y_i Q_i^T)^{k-1} \|$ converges exponentially to 0 as $k \to \infty$. (See Theorem 2, Chap. III.)

DEPENDENCE ON STRUCTURE OF MARKOV CHAIN. The stability properties of the trivial solution are independent of the ordering of the modes. They are, however, dependent on the transient and ergodic set structure of the chain and on the initial distribution of the modes. To demonstrate this dependence, it will be convenient to assume that the transition matrix P of the chain is in the canonic form (4.3) with the E matrices in the form (4.4). In the stability study of a particular system, the availability of P in a canonic form also facilitates the computation of the eigenvalues of $Y_i Q_i^T$, and further, by exhibiting explicitly the transient and ergodic sets aids in the understanding of the problem under consideration. In practice, of course, these advantages may be negated partially by the labor required to reduce the given transition matrix to a canonic form.

Suppose that P is in a canonic form. Also for algebraic convenience, assume that there are two ergodic sets $\mathcal{E}_1$, $\mathcal{E}_2$ and a transient set $\mathcal{T}$. Then Q_i assumes the form

$$Q_i = P_i \otimes I_{n^i} = \begin{bmatrix} E_1 \otimes I_{n^i} & 0 & 0 \\ 0 & E_2 \otimes I_{n^i} & 0 \\ S_1 \otimes I_{n^i} & S_2 \otimes I_{n^i} & T \otimes I_{n^i} \end{bmatrix} \qquad (4.10)$$

Let $U_1, U_2, \ldots, U_{m_1}$, be the modes of $\mathcal{E}_1$, let $V_1, V_2, \ldots, V_{m_2}$ be the modes of $\mathcal{E}_2$, and let $W_1, W_2, \ldots, W_{m_3}$ be the modes of $\mathcal{T}$.

Define

$$Y_i^{(1)} = U_{1[i]} \oplus \cdots \oplus U_{m_1[i]}$$

$$Y_i^{(2)} = V_{1[i]} \oplus \cdots \oplus V_{m_2[i]}$$

$$Y_i^{(3)} = W_{1[i]} \oplus \cdots \oplus W_{m_3[i]}$$

Then, by definition of Y_i,

$$Y_i = Y_i^{(1)} \oplus Y_i^{(2)} \oplus' Y_i^{(3)} \quad ,$$

whence

$$Y_i Q_i^T = \begin{bmatrix} Y_i^{(1)} (E_1^T \otimes I_{n^i}) & 0 & Y_i^{(1)} (S_1^T \otimes I_{n^i}) \\ 0 & Y_i^{(2)} (E_2^T \otimes I_{n^i}) & Y_i^{(2)} (S_2^T \otimes I_{n^i}) \\ 0 & 0 & Y_i^{(3)} (T^T \otimes I_{n^i}) \end{bmatrix} \tag{4.11}$$

The structure of the matrix on the right-hand side of the above equation shows that the eigenvalues of $Y_i^{(1)} (E_1^T \otimes I_{n^i})$, of $Y_i^{(2)} (E_2^T \otimes I_{n^i})$, and of $Y_i^{(3)} (T^T \otimes I_{n^i})$ are the eigenvalues of $Y_i Q_i^T$. The computation of the eigenvalues of $Y_i Q_i^T$ has been reduced to the computation of the eigenvalues of lower-order matrices.

If E_1 is a cyclic set, some simplification in the determination of the eigenvalues of $Y_i^{(1)} (E_1^T \otimes I_{n^i})$ is possible. This is shown in Appendix F.

If the initial distribution $\{a_i\}$ assigns zero probability to all modes not in $\mathcal{E}_1$ (or indeed, in any closed set of modes), then the system behavior must be independent of the values of p_{ij}, X_i, for all i, j $\notin \mathcal{E}_1$. For, $\mathcal{E}_1$ is a <u>closed</u> set; hence, by the choice of the initial distribution, the system can never enter any mode not in $\mathcal{E}_1$. In the case of the ith moment this result can be obtained algebraically by noting that if $a_j = 0$ for all $j \notin \mathcal{E}_1$, then $(Y_i Q_i^T)^k Y_i \mathcal{Q}_i$, and hence $Ex_{k[i]}$, via Eq. (4.8), is for all k independent of p_{ij}, X_i, for all i, $j \notin \mathcal{E}_1$.

In addition to assuming that $\{a_i\}$ assigns zero probability to all modes not in $\mathcal{E}_1$, suppose further that every eigenvalue of $Y_i^{(1)}(E_1^T \otimes I_{n^i})$ is inside the unit circle. Then the trivial solution is asymptotically stable in the ith moment. But by proper choice of $Y_i^{(2)}$, $Y_i^{(3)}$, E_2, T, the matrices $Y_i^{(2)}(E_2^T \otimes I_{n^i})$ and $Y_i^{(3)}(T^T \otimes I_{n^i})$, and hence the matrix $Y_i Q_i^T$, certainly can have eigenvalues outside the unit circle. Hence the converse of the preceding theorem is false; namely, for the trivial solution to be asymptotically stable in the ith moment, it is not necessary that all eigenvalues of $Y_i Q_i^T$ have moduli less than unity. Likewise, if the trivial solution is stable in the ith moment, it is not necessary that no eigenvalue of $Y_i Q_i^T$ lie outside the unit circle and that any eigenvalue on the unit circle have index one.

Suppose now that all eigenvalues of $Y^{(1)}(E_1^T \otimes I_{n^i})$ and $Y^{(3)}(T^T \otimes I_{n^i})$ are inside the unit circle and that the elements of the matrices are such that for some initial distribution the trivial solution is unstable in the ith moment. (This could occur if all elements of all the matrices V_j, $j = 1, \ldots, m_2$ are greater than one.) By choosing an initial distribution which assigns zero probability to all modes in $\mathcal{E}_2$, that is, by choosing $a_{m_1+1}, a_{m_1+2}, \ldots, a_{m_1+m_2} = 0$, the trivial solution becomes stable, in fact, asymptotically stable, in the ith moment. This shows that stability in the ith moment and asymptotic stability in the ith moment depend on the choice of the initial distribution $\{a_i\}$.

It is a fact that with probability one the chain remains in the transient set through only a finite number of transitions.[30] Yet stability in the ith moment and asymptotic stability in the ith moment depend, in general, on p_{ij}, X_i, $i \in \mathcal{T}$. For, consider the simplest possible problem of studying the first moment stability of a first-order system with one absorbing and one transient mode. The P and Y_1 matrices are assumed to be

$$P = \begin{bmatrix} 1 & 0 \\ S & T \end{bmatrix}, \quad T > 0 \tag{4.12}$$

$$Y_1 = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}, \quad X_1, X_2 > 0$$

Hence

$$Y_1 Q_1^T = \begin{bmatrix} X_1 & X_1 S \\ 0 & X_2 T \end{bmatrix}$$

and

$$(Y_1 Q_1^T)^k = \begin{bmatrix} X_1^k & h_k \\ 0 & X_2^k T^k \end{bmatrix}$$

where h_k is a function of X_1, X_2, S, and T. Using Eq. (4.8) to evaluate Ex_k,

$$Ex_k = (a_1 X_1^k + a_2 h_{k-1} X_2 + a_2 X_2^k T^{k-1})\, x_0$$

which clearly depends on S, T, X_2, and $\{a_1, a_2\}$.

The results of the preceding four paragraphs show that if the chain contains more than one ergodic set and possibley a transient set, then it may be possible by changing the initial distribution on the modes to make a stable trivial solution unstable and vice versa. Suppose now that the chain is ergodic, that is, there is only one ergodic set and the transient set is empty. The question arises as to whether (a) the convergence of divergence of $Ex_{k[i]}$ with k depends on the initial distribution of modes; (b) the converse of the preceding theorem is true, namely, is it possible to ensure stability, and asymptotic stability, in the ith moment, by knowing the location and index of the eigenvalues of $Y_iQ_i^T$, if the chain is ergodic. The second question * is not answered here. A partial answer to the first question is given by the following example which shows that the value of $Ex_{k[i]}$ for large k can depend markedly on the initial distribution.

The example is as follows: The system is of order two, has two modes, and $t_k - t_{k-1} = 1$ for all $k \geq 1$. The random matrix A_k can assume the two values

$$\begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}, \quad \alpha \text{ fixed}, \quad \beta = 2\pi \text{ or } \pi.$$

Hence the two respective values of $\Phi_k = \exp[A_k(t_k - t_{k-1})]$ are

$$X_1 = e^{\alpha} I_2 \quad \text{and} \quad X_2 = -e^{\alpha} I_2$$

where I_n is the nth-order identity matrix. Let the transition matrix of the chain be

$$P = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}, \quad 0 < p < 1$$

Then

$$Y_1 Q_1^T = \begin{bmatrix} e^{\alpha} I_2 & 0 \\ 0 & -e^{\alpha} I_2 \end{bmatrix} (P^T \otimes I_2) \tag{4.13}$$

$$= (e^{\alpha} L \otimes I_2)(P^T \otimes I_2) = (e^{\alpha} L P^T) \otimes I_2 \tag{4.14}$$

*The answer to the second question is in the affirmative if the chain is regular and all elements of X_j are nonnegative. This case will be considered in more detail in a later report.

where

$$L = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Equation (4.14) follows from (4.13) by the commutative property (3.11) of the Kronecker product. Using this commutative property again,

$$(Y_1Q_1^T)^2 = \left((e^{a}LP^T) \otimes I_2\right)\left((e^{a}LP^T) \otimes I_2\right)$$

$$= (e^{a}LP^T)^2 \otimes I_2 = e^{2a}(2p-1)I_2 \otimes I_2$$

It follows that

$$(Y_1Q_1^T)^{2k} = e^{2ka}(2p-1)^k I_4$$

whence by Eq. (4.8)

$$E_{2k+1} = e^{(2k+1)a}(2p-1)^k(a_1-a_2)x_0 \tag{4.15}$$

Evaluating Ex_{2k+1} in a like fashion,

$$Ex_{2k} = e^{2ka}(2p-1)^k x_0 \tag{4.16}$$

Hence if j is odd, Ex_j depends on the initial distribution $\{a_1, a_2\}$; if j is even, it does not. Equations (4.15) and (4.16) show that the trivial solution is stable in the first moment if

$$\left| e^{a}\sqrt{2p-1} \right| \leq 1 \tag{4.17}$$

(Note that $\pm e^{a}\sqrt{2p-1}$ are the eigenvalues of $Y_1Q_1^T$ which checks with the results of Theorem 1.) Equations (4.15) and (4.16) also show that (4.17) is, moreover, a necessary condition for first moment stability. Likewise, (4.17) changed to a strict inequality gives the necessary and sufficient condition for asymptotic stability in the ith moment. In exactly the same fashion it can be demonstrated that the trivial solution is stable (asymptotically stable) in the ith moment if and only if $|e^{i\alpha}\sqrt{2p-1}| \leq 1$ $(e^{i\alpha}\sqrt{2p-1} < 1)$.

4. 3 Almost Sure Asymptotic Stability

As in the independent case, there exists the following

THEOREM 2: The trivial solution is equiasymptotically stable almost surely if for some even i, all eigenvalues of $Y_iQ_i^T$ are inside the unit circle. (Compare with Theorem 3, Chap. III.)

PROOF: The proof is virtually identical to that of Theorem 3 of Chap. III.

By Eq. (4.8),

$$Ex_{k[i]} = \mathcal{J}_i(Y_iQ_i^T)^{k-1}Y_i\mathcal{A}_ix_{0[i]}$$

Summing on k,

$$\sum_{k=1}^{\infty} Ex_{k[i]} = \mathcal{J}_i \sum_{k=1}^{\infty} (Y_iQ_i^T)^{k-1}Y_i\mathcal{A}_ix_{0[i]}$$

$$= \mathcal{J}_i(I_{mn^i} - Y_iQ_i^T)^{-1}Y_i\mathcal{A}_ix_{0[i]}$$

because of assumption on the eigenvalues of $Y_iQ_i^T$.

The remainder of the proof is exactly the same as that of Theorem 3 of Chap.III.

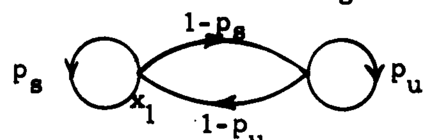
The converse of the above theorem is false as is shown by Example 1 of Chap.III.

4. 4 Example

Consider the same system as in Example 1 of Chap.III but with a Markov structure:

A_k can assume two possible values, $b_1 \leq 0$ and $b_2 \geq 0$; $t_k - t_{k-1} = T$

for all k. Hence Φ_k can assume the two possible values $X_1 = e^{b_1 T}$ and $X_2 = e^{b_2 T}$. The transition diagram for the Markov chain is



with corresponding transition matrix

$$P = \begin{bmatrix} p_s & 1-p_s \\ 1-p_u & p_u \end{bmatrix}$$

Since the system is of order 1, $Q_i = P$, and

$$Y_i = \begin{bmatrix} X_1^i & 0 \\ 0 & X_2^i \end{bmatrix}$$

Since all entries of the matrix $Y_i Q_i^T$ are positive, it can be shown that a necessary <u>and</u> sufficient condition for the asymptotic stability in the ith moment is that every eigenvalue of $Y_i Q_i^T$ lies inside the unit circle. Solving for the eigenvalues

$$| Y_i Q_i^T - \lambda I_2 | = \lambda^2 - \lambda (p_s X_1^i + p_u X_2^i) + (1 - p_s - p_u) X_1^i X_2^i = 0$$

The Schur-Cohn criterion can be used on this polynomial in λ to obtain Figs. 4-6 which show the regions of stability in the p_s-p_u plane. The line segment labeled "independent" in the figures is the set of points (p_s, p_u) such that the Markov chain becomes an independent, identically distributed, random sequence.

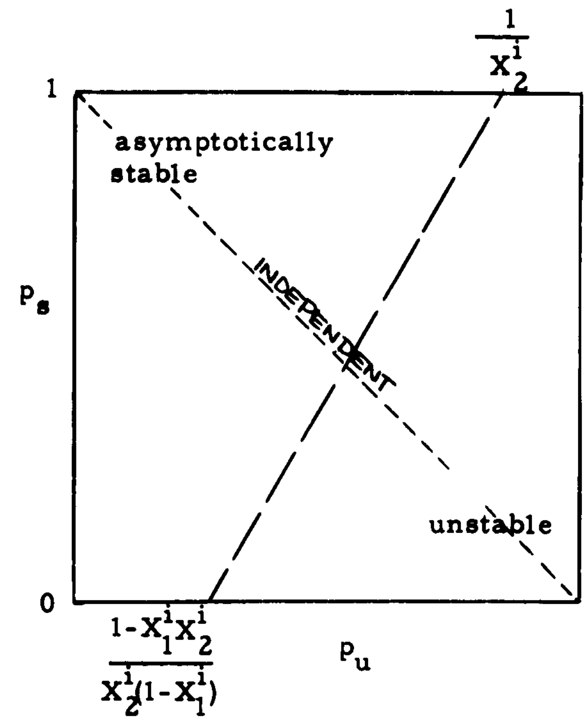


Figure 4

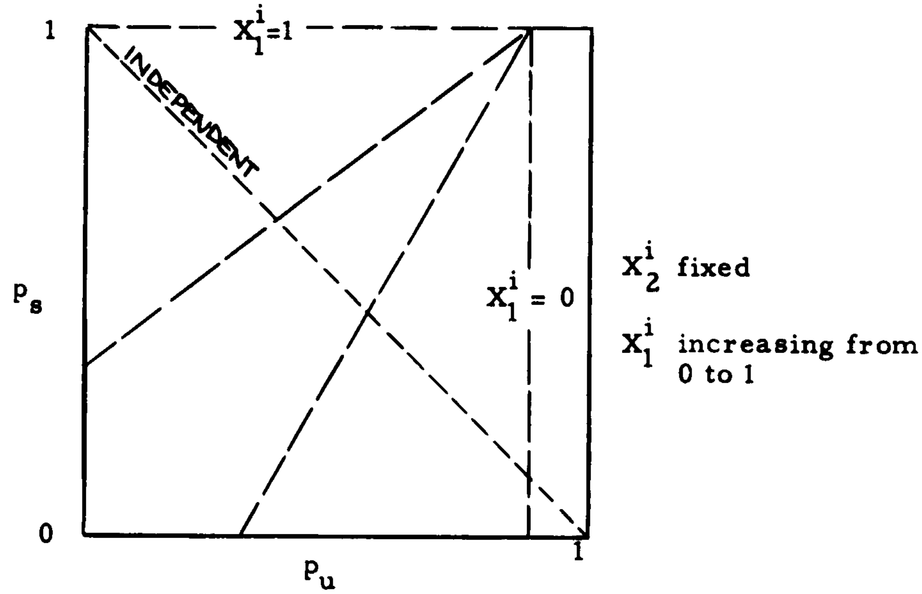


Figure 5

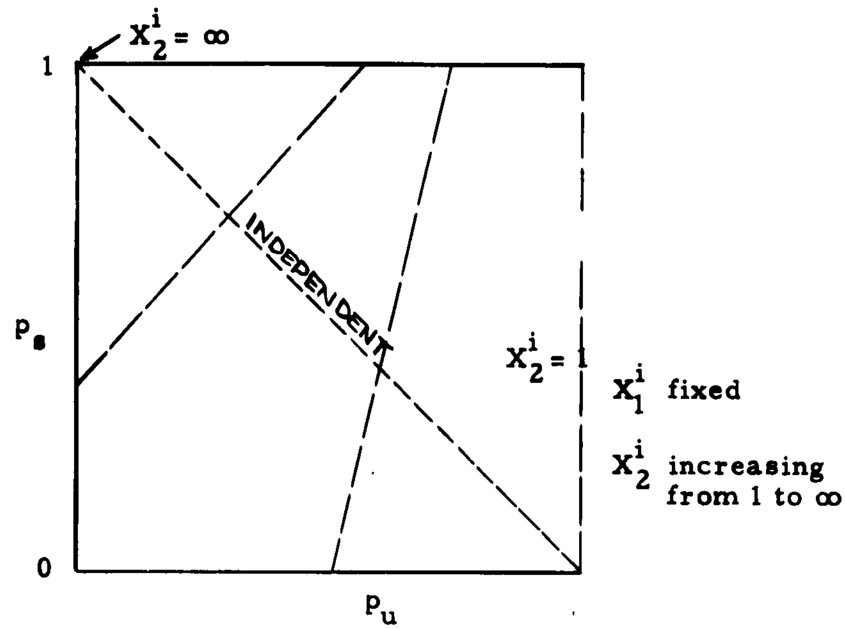


Figure 6

4.5 Some General Comments

In both Chaps. III and IV, the fact that the matrices Φ_k are of the form exp[B] has not been made use of. Hence the results pertain not only to piecewise constant differential equations but also to all random linear difference equations $x_k = \Phi_k x_{k-1}$, where $\{\Phi_k\}$ is an independent, identically distributed, random process, or more generally, a finite Markov chain.

Such difference equations arise in the study of randomly sampled linear systems, whose second moment stability has been investigated by Kalman.[6] Kalman considers the case where the successive sampling intervals are independent, identically distributed, random variables. This corresponds to the situation when $\{\Phi_k\}$ is an independent, identically distributed, random process. Kalman's method consists of taking the expectation of the scalar product,

$$E < x_k, x_k > = E < \Phi_k x_{k-1}, \Phi_k x_{k-1} >$$

$$= E < \Phi_k \cdots \Phi_1 x_0, \Phi_k \cdots \Phi_1 x_0 >$$

$$= E(x_0^T \Phi_1^T \cdots \Phi_k^T \Phi_k \cdots \Phi_1 x_0).$$

which leads him to the use of power product matrices. The method is not as straightforward as the one presented here and suffers from the limitation that it cannot be directly extended to the study of moments of order > 2, and to other statistical structures such as Markov chains.

All of the material presented in this section can be extended to the case where $\{\Phi_k\}$ is a multiple-order Markov chain. This extension will be included in a later report.

V. THE FORCED LINEAR SYSTEM

Summary

For random linear systems, a theorem is proven which gives sufficient conditions on the input and on the solutions of the unforced system, in order that solutions of the forced system be bounded in the mean norm. The results of the theorem are immediately applicable to the piecewise constant, linear system considered in Chaps. III and IV.

The preceding chapters consider the stability, with respect to changes in initial conditions, of a fixed solution function. Hence, the input to the system is necessarily assumed to be fixed, usually at zero. In practice, however, it is the behavior of the system in the presence of any input, belonging to some class of possible inputs, that is of interest.

For deterministic linear systems a common definition of stability is: the output to every bounded input is bounded. Because of linearity this definition of stability is equivalent[19, 33] to any of the following properties of the unforced linear system: (i) the trivial solution is uniformly asymptotically stable; (ii) the trivial solution is exponentially asymptotically stable; (iii) $\int_{t_0}^{t} \| W(t, \tau) \| \, d\tau$ is bounded in (t, t_0), $t \geq t_0$, where $W(t, \tau)$ is the fundamental matrix.

The following theorem partially generalizes some of these results to the random case:

THEOREM 1: Consider the vector differential equation

$$\dot{x} = A(t)x + b(t), \qquad t \geq 0 \tag{5.1}$$

where both $\{A(t)\}$ and $\{b(t)\}$ are random processes and $\sup_t \|A(t)\|$ is a. s. bounded. For almost every ω, assume that A and b are sufficiently smooth[35] so that solutions of (5.1) exist and are unique. Let $W(t, \tau)$, $t, \tau > 0$, be the fundamental matrix of the homogeneous equation ($b \equiv 0$) which is normalized to $W(t,t) = I$, $t \geq 0$. Denote the solution of (5.1) by $g_b(t, x_0, t_0)$. Now consider the following propositions (M_i are finite positive constants independent of (t, t_0)):

(a) The trivial solution of the homogeneous equation is exponentially asymptotically stable in the mean norm; or equivalently, there exist two constants K, $\nu > 0$ such that

$$E\| W(t, t_0)\| < K \exp[-\nu(t-t_0)], \qquad t \geq t_0.$$

(b) $$\int_{t_0}^{t} E\| W(t, \tau)\| \, d\tau < M_1, \qquad t \geq t_0.$$

(c) If $E\| b(t)\| < M_2$, and $\{b(t)\}$ is statistically independent of

Proposition (a) states that the unforced system behaves, in the mean norm sense, like an unforced, constant coefficient linear system all of whose characteristic roots have negative real parts.

Propositions (c) and (c') are stochastic equivalents of the oft-heard definition of stability for deterministic linear systems: to every bounded input there corresponds a bounded output.

The theorem gives a partial justification for considering, in stability studies of random linear systems, the simpler unforced system instead of the whole family of forced systems. This simplification, of recognized value in the study of deterministic linear systems, is perhaps of even greater value for random linear systems because of the more complicated nature of the problem.

Before proving the theorem, consider the following

LEMMA: Proposition (b) of Theorem 1 implies that

$\| W(t, t_0) \| < M_6, \; t \geq t_0 \geq 0.$

PROOF: The proof given below follows the one given by Kalman[19] for the deterministic case.

The fundamental matrix satisfies the homogeneous differential equation:

$$\frac{d}{dt} W(t, \tau) = A(t) W(t, \tau)$$

Hence

$$\begin{aligned} \frac{d}{dt} W^{-1}(t, \tau) &= -W^{-1}(t, \tau) \frac{d}{dt} W(t, \tau) W^{-1}(t, \tau) \\ &= -W^{-1}(t, \tau) A(t) W(t, \tau) W^{-1}(t, \tau) \\ &= -W^{-1}(t, \tau) A(t) \end{aligned}$$

On interchanging t and τ,

$$\frac{d}{d\tau} W^{-1}(\tau, t) = -W^{-1}(\tau, t) A(\tau)$$

from which it follows that

$$\frac{d}{d\tau} W(t, \tau) = -W(t, \tau)A(\tau)$$

since $W^{-1}(\tau, t) = W(t, \tau)$. Integrate with respect to τ, use the fact that $W(t, t) = I$, and then take norms:

$$W(t, t_0) - I = \int_{t_0}^{t} -W(t, \tau)A(\tau)d\tau$$

$$\| W(t, t_0) - I\| \leq \int_{t_0}^{t} \| W(t, \tau)\| \; \| A(\tau)\| \; d\tau$$

$$\leq [\text{a. s. } \sup(\sup_{t\geq 0} \| A(t)\|)] \int_{t_0}^{t} \| W(t, \tau)\| \; d\tau .$$

Now $\| W(t, t_0)\| \leq \| W(t, t_0) - I\| + \| I\|$. Hence, upon taking expectations,

$$E\| W(t, t_0)\| \leq [\text{a. s. } \sup (\sup_{t\geq 0} \| A(t)\|)] \, E \int_{t_0}^{t} \| W(t, \tau)\| \; d\tau + E\| I\|$$

$$< [\text{a. s. } \sup (\sup_{t\geq 0} \| A(t)\|)] \, M_1 + \| I\|$$

Hence $E\| W(t, t_0)\|$ is bounded in (t, t_0), $t \geq t_0$.

PROOF of Theorem 1: To show (a) $\Longrightarrow$ (b).

$$\int_{t_0}^{t} E\| W(t, \tau)\| \; d\tau < \int_{t_0}^{t} Ke^{-\nu (t-\tau)} d\tau < \frac{K}{\nu}$$

To show (b) $\Longrightarrow$ (c) . The solution of (5.1) can be expressed as

$$g_b(t, x_0, t_0) = W(t, t_0)x_0 + \int_{t_0}^{t} W(t, \tau)b(\tau) \, d\tau$$

which upon taking norms becomes

$$\| g_b(t, x_0, t_0) \| \leq \| W(t, t_0) \| \cdot \| x_0 \| + \int_{t_0}^{t} \| W(t, \tau) \| \; \| b(\tau) \| \; d\tau \tag{5.2}$$

Since $\{A(t)\}$ and $\{b(t)\}$ are independent,

$$E \| g_b(t, x_0, t_0) \| \leq E \| W(t, t_0) \| \cdot \| x_0 \| + \int_{t_0}^{t} E \| W(t, \tau) \| \cdot E \| b(\tau) \| \, d\tau .$$

By the preceding Lemma, $E \| W(t, t_0) \| < M_6$; and by hypothesis, $E \| b(\tau) \| < M_2$. Hence

$$E \| g_b(t, x_0, t_0) \| < M_6 \| x_0 \| + M_2 \int_{t_0}^{t} E \| W(t, \tau) \| \, d\tau$$

$$< M_6 \| x_0 \| + M_2 M_1$$

$$= M_3(M_1, \| x_0 \|)$$

To show (b)$\Longrightarrow$(c'). From inequality (5.2)

$$\| g_b(t, x_0, t_0) \| \leq \| W(t, t_0) \| \cdot \| x_0 \| + [\text{a.s. } \sup (\sup_{t \geq 0} \| b(t) \|)] \int_{t_0}^{t} \| W(t, \tau) \| \, d\tau$$

Therefore,

$$E \| g_b(t, x_0, t_0) \| < E \| W(t, t_0) \| \cdot \| x_0 \| + M_4 \int_{t_0}^{t} E \| W(t, \tau) \| \, d\tau$$

whence by the preceding Lemma, and by (b)

$$E \| g_b(t, x_0, t_0) \| < M_6 \| x_0 \| + M_4 M_1$$

$$= M_5(M_1, \| x_0 \|)$$

The results of the preceding theorem are immediately applicable to the random linear differential system considered in Chaps. III and IV. The vector differential equation now becomes

$$\dot{x} = A_k x + b(t), \ t_{k-1} \leq t < t_k, \ k = 1, 2, \dots \tag{5.3}$$

Suppose that the trivial solution of the homogeneous system is exponentially asymptotically stable in the ith moment for some even i. Then, as shown in Chap. II, it is exponentially asymptotically stable in the mean norm. Hence, by the theorem just proven, it follows that

> If for the differential system (5.3), the trivial solution of the homogeneous system is exponentially asymptotically stable in the ith moment for some even i, then $E\| g_b(t, x_0, t_0)\|$ is bounded uniformly in (t, t_0) for all bounded x_0, provided either
>
> > $\{b(t)\}$ is statistically independent of $\{A_k(t_k - t_{k-1})\}$ and $E\| b(t)\|$ is bounded uniformly in t,
>
> or
>
> > $\{b(t)\}$ is a. s. bounded.

APPENDIX A

KRONECKER AND POWER PRODUCTS OF MATRICES*

In the definitions and identities that follow, except for finiteness, no assumptions about the dimensions of the matrices are made unless indicated explicitly, or implied by the use of operations such as inverse, trace, determinant, the conventional sum A + B and product AB, etc.

KRONECKER PRODUCTS

If A and B are matrices of order α and β, respectively, of the transformations

*For a concise statement of the properties of these products, see MacDuffee,[22] Chap. VII.

$$\xi_i' = \sum_{j=1}^{\alpha} a_{ij}\xi_j \ , \qquad \eta_i' = \sum_{j=1}^{\beta} b_{ij}\eta_j \ , \tag{A.1}$$

the product vector (column)

$$(\xi_1'\eta_1' , \xi_1'\eta_2' , \ldots , \xi_1'\eta_\beta' , \xi_2'\eta_1' , \xi_2'\eta_2' , \ldots , \xi_2'\eta_\beta' , \ldots , \xi_\alpha'\eta_\beta') \tag{A. 2}$$

is a linear transform of the corresponding unprimed product vector (column)

$$(\xi_1\eta_1 , \xi_1\eta_2 , \ldots , \xi_1\eta_\beta , \xi_2\eta_1 , \xi_2\eta_2 , \ldots , \xi_2\eta_\beta , \ldots , \xi_\alpha\eta_\beta) \tag{A. 3}$$

where the linear transformation is given by the Kronecker (or direct) product matrix:

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1\alpha}B \\ a_{21}B & a_{22}B & \cdots & a_{2\alpha}B \\ \vdots & \vdots & & \\ a_{\alpha 1}B & a_{\alpha 2}B & \cdots & a_{\alpha\alpha}B \end{bmatrix}$$

DEFINITION: Given an $\alpha_1 \times \alpha_2$ matrix $A = (a_{ij})$ and a $\beta_1 \times \beta_2$ matrix $B = (b_{ij})$, the Kronecker (or direct) product of A by B is the $\alpha_1\beta_1 \times \alpha_2\beta_2$ matrix

$$A \otimes B = (a_{ij}B) = \begin{bmatrix} a_{11}B & \cdots & a_{1\alpha_2}B \\ \vdots & & \vdots \\ a_{\alpha_1 1}B & \cdots & a_{\alpha_1\alpha_2}B \end{bmatrix}$$

The product is associative

$$(A \otimes B) \otimes C = A \otimes (B \otimes C)$$

from which follows the definition of the ith self-Kronecker product:

$$A_{[i]} = A \otimes \ldots \otimes A$$

to i factors. It obeys the relation

$$AB_{[i]} = A_{[i]}B_{[i]} \qquad \text{(A.4)}$$

In fact, the more general identity

$$A_1B_1 \otimes A_2B_2 \otimes \ldots \otimes A_iB_i = (A_1 \otimes A_2 \otimes \ldots \otimes A_i)(B_1 \otimes B_2 \otimes \ldots \otimes B_i)$$

holds, from which it follows that if

$$Ax_i = \lambda_i x_i \qquad By_j = \mu_j y_j$$

then

$$(A \otimes B)(x_i \otimes y_i) = Ax_i \otimes By_j = \lambda_i x_i \otimes \mu_j y_j = \lambda_i \mu_j (x_i \otimes y_j), \qquad \text{(A. 5)}$$

that is, if λ_i and μ_j are eigenvalues of A and B, respectively, with corresponding eigenvectors x_i and y_j, then $A \otimes B$ has the eigenvalue $\lambda_i \mu_j$ with the associated eigenvector $x_i \otimes y_j$.

Most of the following identities are readily verified:

(a) $(A + B) \otimes C = A \otimes C + B \otimes C$

(b) $(A \otimes B)^T = A^T \otimes B^T$, T denotes transpose.

If A and B are matrices of orders α and β, respectively,

(c) trace $(A \otimes B) = (\text{trace } A)(\text{trace } B)$

(d) $\det(A \otimes B) = \det(A^\beta)\det(B^\alpha)$

(e) $A \otimes B = (A \otimes I_\beta)(I_\alpha \otimes B)$, I_α = identity matrix of order α

(f) $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$

(g) $(L_1 \otimes M_1)(A \otimes B)(L_2 \otimes M_2) = A_1 \otimes B_1$,

where $L_1AL_2 = A_1$ and $M_1BM_2 = B_1$

(h) $(L \otimes M)^{-1}(A \otimes B)(L \otimes M) = A_1 \otimes B_1$,

where $L^{-1}AL = A_1$ and $M^{-1}BM = B_1$

POWER PRODUCTS

In (A.1), let the two transformations be identical. Then the $(1/2)\alpha(\alpha+1)$ distinct products

$$\xi_1'^2, \; \xi_1' \xi_2', \ldots, \xi_1' \xi_\alpha', \; \xi_2'^2, \; \xi_2' \xi_3', \ldots, \xi_\alpha'^2$$

are related to the corresponding unprimed products by a transformation whose matrix $A_{(2)}$ is called the second power matrix of A. In symbols

$$x' \cdot\!\times\; x' = (A \cdot\!\times\; A)(x \cdot\!\times\; x)$$

$$x'_{(2)} = A_{(2)} x_{(2)}$$

where $x = (\xi_1, \ldots, \xi_\alpha)$ and $x' = (\xi_1', \ldots, \xi_\alpha')$.

The ith power product of x is similarly defined as the vector $x_{(i)}$ having as components the $\binom{\alpha+i-1}{i}$ distinct ith-degree products of components of x, arranged in the lexicographic order. $x_{(i)}$ and $x'_{(i)}$ are again linearly related:

$$x'_{(i)} = A_{(i)} x_i$$

where $A_{(i)}$ is the matrix of the linear transformation.

The lexicographic ordering, of the components of x, is merely a convenient one. If some other ordering is used, the rows and columns of $A_{(i)}$ will be correspondingly permuted.

The Kronecker and power products have several similar properties. In particular, the following theorems hold:

(a) $AB_{(i)} = A_{(i)} B_{(i)}$

If A is of order α, then

(b) $\det A_{(i)} = \det(A^j)$, $\qquad j = \binom{i+\alpha-1}{\alpha}$

and

(c) $A_{(i)}$ has as eigenvalues the $\binom{\alpha + i - 1}{i}$ products of the ith degree of the eigenvalues of A.

APPENDIX B

A MODIFIED HÖLDER INEQUALITY

Let Z_j, $j = 1, 2, \ldots, n$, be scalar random variables. Then

$$E \prod_j | Z_j^{i_j} | \leq \prod_j (E| Z_j^i |)^{i_j/i} \tag{B.1}$$

$$\sum_j i_j = i, \qquad i_j > 0 \text{ for all } j$$

PROOF: In the case of two random variables X_1, X_2, there exists the well-known Hölder Inequality[21]

$$E| X_1 X_2| \leq (E| X_1^{r_1}|)^{1/r_1} (E| X_2^{r_2}|)^{1/r_2}, \tag{B. 2}$$

$$\frac{1}{r_1} + \frac{1}{r_2} = 1, \qquad r_1, r_2 > 0$$

Under the substitutions $X_1 = U^j$, $X_2 = V^k$, $1/r_1 = j/(j+k)$, $1/r_2 = 1 - 1/r_1 = k/(j+k)$, (B. 2) becomes

$$E| U^j V^k| \leq (E| U^{j+k}|)^{j/(j+k)} (E| V^{j+k}|)^{j/(j+k)}, \tag{B.3}$$

$$j, k > 0$$

Let $V^k = Z_1^{k_1} Z_2^{k_2}$, $k_1 + k_2 = k$, $k_1, k_2 > 0$. Then

$$V^{j+k} = Z_1^{\frac{k_1}{k}(j+k)} \quad Z_2^{\frac{k_2}{k}(j+k)} \quad ;$$

note that the sum of the exponents of Z_1 and Z_2 is (j+k).
Upon substituting for V, (B.3) becomes

$$\left(E|U^j Z_1^{k_1} Z_2^{k_2}|\right) \leq \left(E|U^{j+k}|\right)^{\frac{j}{j+k}} \left(E|Z_1^{\frac{k_1}{k}(j+k)} Z_2^{\frac{k_2}{k}(j+k)}|\right)^{\frac{j}{(j+k)}}$$

Now apply inequality (B.3) to the extreme right-hand term in the above inequality to obtain

$$\left(E|U^j Z_1^{k_1} Z_2^{k_2}|\right) \leq \left(E|U^{j+k}|\right)^{\frac{j}{j+k}} \left(E|Z_1^{j+k}|\right)^{\frac{k_1}{j+k}} \left(E|Z_2^{j+k}|\right)^{\frac{k_2}{j+k}}$$

which can be successively generalized to (B. 1).

APPENDIX C

SOME THEOREMS ON ALMOST SURE AND ALMOST UNIFORM-IN-ω TYPES OF STABILITY

Theorems 1, 3 and 4, relating almost sure and almost uniform-in-ω types of stability are proven. The proofs are modeled after those of Egoroff's Theorem (for finite measure, convergence almost everywhere implies almost uniform convergence) and its "converse" (almost uniform convergence implies convergence almost everywhere).[34]

THEOREM 1: The trivial solution of the differential system (2. 2) is stable almost surely (A) if, and (B) only if, it is stable almost uniformly-in-ω.

PROOF: (A) By hypothesis, given any positive integer n and any t_0, there exists an ω-set $B_n(t_0)$ with $PB_n > 1 - 1/n$, and given any $\epsilon > 0$, there is a corresponding $\delta(\epsilon, n, t_0) > 0$ such that $\|x_0\| < \delta$ implies that $\|g(t, x_0, t_0; \omega)\| < \epsilon$ for all $t \geq t_0$ and all $\omega \in B_n$.

Let $S = \bigcup_{n=1}^{\infty} B_n$. Then for every n, $S \supset B_n$ and

$PS \geq PB_n > 1 - 1/n$. Hence $PS = 1$.

To show that the trivial solution is stable for all realizations corresponding to the points of $S = S(t_0)$, choose any $\omega' \in S$. Then $\omega' \in B_k$ for some $k = k(\omega')$ whence for any t_0 and any $\epsilon > 0$, $\| x_0 \| < \delta(\epsilon, k(\omega'), t_0)$ implies that $\| g(t, x_0, t_0; \omega') \| < \epsilon$ for all $t \geq t_0$.

Given any other initial time t_0', $S(t_0')$ can be taken to be $S(t_0)$. For, if the trivial solution of the deterministic system corresponding to $\omega \in S(t_0)$ is stable at t_0, it is stable at t_0'; that is, given any $\epsilon > 0$, there is a $\delta(\epsilon, t_0', \omega) > 0$ such that $\| x_0 \| < \delta(\epsilon, t_0', \omega)$ implies $\| g(t, x_0, t_0'; \omega) \| < \epsilon$ for all $t \geq t_0'$. Hence there is a set S of unit probability on which the trivial solution is stable at every t_0. Therefore the trivial solution is stable almost surely.

(B) by hypothesis, there exists an ω-set S of unit probability, and to every $\omega \in S$, every positive integer m, and every t_0, there corresponds a positive integer $\bar{n}(m, \omega, t_0)$ such that $\| x_0 \| < 1/\bar{n}$ implies $\| g(t, x_0, t_0; \omega) \| < 1/m$ for all $t \geq t_0$.

Given any t_0, let

$$E_n^m(t_0) = E_n^m = \{\omega : \| x_0 \| < \frac{1}{n} \Longrightarrow \| g(t, x_0, t_0; \omega) \| < \frac{1}{m}$$

$$\text{for all } t \geq t_0\}, \quad m, \ n = 1, 2, \ldots$$

If $\omega \in S$, then for every m there is an $n_1(m, \omega, t_0) > 0$ such that $\omega \in E_{n_1}^m$. Therefore $\bigcup_n E_n^m \supset S$, whence $P \bigcup_n E_n^m = 1$. Since $E_1^m \subset E_2^m \subset E_3^m \subset \ldots$, $\lim_n PE_n^m = 1$, and given any $\eta > 0$ there exists a smallest integer $N(m, \eta, t_0)$ such that $PE_N^m > 1 - \eta/2^m$.

Define $B = \bigcap_m E_N^m$. Note that B depends on η and t_0 only.

$$PB^c = P \bigcup_m (E_N^m)^c \leq \sum_m P(E_N^m)^c \leq \sum_m \frac{\eta}{2^m} = \eta .$$

Hence $PB > 1 - \eta$. Moreover, for the t_0 chosen above, if $\omega' \in B$, then

$\omega' \varepsilon E_N^m$ for all m so that $\| x_0 \| < 1/N(m, \eta, t_0)$ implies $\| g(t, x_0, t_0; \omega') \| < 1/m$ for all m and for all $t \geq t_0$. Therefore the trivial solution is stable almost uniformly-in-ω .

THEOREM 3: The trivial solution of (2. 2) is quasi-asymptotically stable almost surely if it is quasi-asymptotically stable almost uniformly-in-ω with the set B (see definition (iii-f), Sec. II) independent of x_0.

PROOF: By hypothesis, given any integer $n > 0$ and any t_0, there is a $\delta(n, t_0) > 0$ and an ω-set $B_n(t_0)$ with $PB_n > 1 - 1/n$ such that $\| x_0 \| < \delta$ implies that $g(t, x_0, t_0; \omega) \longrightarrow 0$ as $t \longrightarrow \infty$ uniformly on B_n.

Let $S(t_0) = \bigcup_n B_n(t_0)$. Then for all n, $S(t_0) \supset B_n(t_0)$ and $PS(t_0) \geq PB_n(t_0) > 1 - 1/n$. Hence $PS(t_0) = 1$.

Given any $\omega \varepsilon S(t_0)$, ω is an element of $B_k(t_0)$ for some $k = k(\omega)$. Hence at t_0, $\| x_0 \| < \delta(k(\omega), t_0)$ implies $g(t, x_0, t_0; \omega) \rightarrow 0$ as $t \rightarrow \infty$; i. e., the trivial solution is quasi-asymptotically stable almost surely at t_0. As in the proof of Theorem 1, S can be taken to be independent of t_0. This completes the proof.

THEOREM 4: The trivial solution of (2. 2) is quasi-asymptotically stable almost uniformly-in-ω if it is quasi-asymptotically stable almost surely.

PROOF: By hypothesis, there exists an ω-set S of unit probability and to every $\omega \varepsilon S$ and any given t_0, there corresponds a positive integer $\bar{n}(\omega, t_0)$ such that $\| x_0 \| < 1/\bar{n}$ implies $g(t, x_0, t_0; \omega) \rightarrow 0$ as $t \rightarrow \infty$.

Given any t_0, let

$$F_n(t_0) = F_n = \{ \omega : \| x_0 \| < \frac{1}{n} \Longrightarrow g(t, x_0, t_0; \omega) \rightarrow 0 \text{ as } t \rightarrow \infty \}, \quad n = 1, 2, \ldots$$

Pick $\omega \varepsilon S$. Then $\omega \varepsilon F_{\bar{n}(\omega, t_0)}$ and hence $\omega \varepsilon \bigcup_n F_n$. Therefore $\bigcup_n F_n \supset S$ whence $P \bigcup_n F_n = 1$. Since $F_1 \subset F_2 \subset F_3 \ldots$, $\lim_n PF_n = 1$.

Hence, given any $\eta > 0$ there is a smallest integer $N(\eta, t_0)$ such that

$PF_N > 1 - \eta/2$.

For every x_0, $\|x_0\| < 1/N$, define

$$E_k^m(x_0, t_0) = E_k^m = \{\omega : \| g(t, x_0, t_0; \omega)\| < \frac{1}{m} \text{ for all } t \geq t_0 + k\}$$

$$m,\ k = 1,\ 2,\ \ldots$$

If $\omega \varepsilon F_N$, then for every m there is a $k_1(m, \omega, x_0, t_0)$ such that $\omega \ \varepsilon\ E_{k_1}^m$. Therefore $\bigcup_k E_k^m \supset F_N$, whence $P \bigcup_k E_k^m > 1 - \eta/2$. Since $E_1^m \subset E_2^m \subset E_3^m \subset \ldots\ldots \lim_k PE_k^m > 1 - \eta/2$ and there exists a smallest integer $K(m, \eta, x_0, t_0) > 0$ such that $\lim_k PE_k^m - PE_K^m < (\eta/2)/2^m$ from which it follows that $PE_K^m > 1 - (\eta/2)(1 + 1/2^m)$.

Define $B = \bigcap_m E_K^m$ and note that $B = B(\eta, x_0, t_0)$.

$$PB^c = P \bigcup_m (E_K^m)^c \leq \sum_m P(E_K^m)^c \leq \sum_m \frac{\eta}{2}(1 + \frac{1}{2^m}) = \eta$$

Hence $PB > 1 - \eta$. Moreover, if $\omega' \varepsilon B$, then $\omega' \varepsilon E_K^m$ for every m so that $\|x_0\| < 1/N(\eta, t_0)$ implies $\|g(t, x_0, t_0; \omega')\| < 1/m$ for all m and for all $t \geq t_0 + K(m, \eta, x_0, t_0)$. Therefore the trivial solution is quasi-asymptotically stable almost uniformly-in- ω .

APPENDIX D

$E(\Phi_k \Phi_{k-1} \cdots \Phi_1)$ FOR MATRIX VALUED MARKOV CHAIN $\{\Phi_k\}$

Given a finite Markov chain with initial probability vector $(a_1, \ldots, a_m)$ and an m x m constant matrix $P = (p_{ij})$ of transition probabilities, associate with each state i an n x n matrix, X_i. Let Φ_k be the random matrix occurring at time k, k = 1, 2, ... The expected value of the product matrix $\Phi_k \Phi_{k-1} \cdots \Phi_1$ is derived here.

For notational and algebraic simplicity consider a four-fold product. Then

$$E\,\Phi_4\,\Phi_3\,\Phi_2\,\Phi_1 = \sum_{i,j,k,\ell=1}^{m} X_\ell X_k X_j X_i a_i p_{ij} p_{jk} p_{k\ell}$$

$$= \sum_{i,k,\ell} X_\ell X_k \left(\sum X_j p_{ij} p_{jk} \right) X_i a_i p_{k\ell} \qquad \text{(D.1)}$$

Let I be the identity matrix of order n and define Q as the Kronecker product

$$Q = P \otimes I$$

and Y as the direct sum

$$Y = X_1 \oplus X_2 \oplus \dots \oplus X_m$$

Then

$$U = Q^T Y Q^T \qquad \text{(D. 2)}$$

$$= \begin{bmatrix} p_{11}I & p_{21}I & \cdots \\ p_{12}I & p_{22}I & \cdots \\ \vdots & \vdots & \cdots \end{bmatrix} \begin{bmatrix} X_1 & 0 & \cdots \\ 0 & X_2 & \cdots \\ \vdots & \vdots & \cdots \end{bmatrix} \begin{bmatrix} p_{11}I & p_{21}I & \cdots \\ p_{12}I & p_{22}I & \cdots \\ \vdots & \vdots & \cdots \end{bmatrix}$$

$$= \begin{bmatrix} p_{11}X_1p_{11} + p_{21}X_2p_{12} + \cdots & p_{11}X_1p_{21} + p_{21}X_2p_{22} + \cdots & \cdots \\ p_{12}X_1p_{11} + p_{22}X_2p_{12} + \cdots & p_{12}X_1p_{21} + p_{22}X_2p_{22} + \cdots & \cdots \\ \vdots & \vdots & \cdots \end{bmatrix}$$

$$= [U_{ik}] \qquad \text{(D.3)}$$

where U_{ik} is the matrix entry in row i and column k. The parenthesized term on the right side of Eq. (D.1) can be identified as U_{ki} so that (D.1) becomes

$$E\Phi_4\Phi_3\Phi_2\Phi_1 = \sum_{i,\ell} X_\ell \left(\sum_k X_k U_{ki} p_{k\ell}\right) X_i a_i \qquad (D.4)$$

By steps similar to (D. 2) through (D. 3),

$$\sum_k X_k U_{ki} p_{k\ell} = V_{\ell i}$$

where

$$V = [V_{i\ell}] = Q^T Y U \qquad (D.5)$$

Equation (D. 4) now reduces to

$$E\Phi_4\Phi_3\Phi_2\Phi_1 = \sum_\ell X_\ell \sum_i V_{\ell i} X_i a_i \qquad (D.6)$$

Let $\mathcal{a}$ be the Kronecker product of the initial probability vector with the identity:

$$\mathcal{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \otimes I$$

Let $W_\ell = \sum_i V_{\ell i} X_i a_i$ and let W be the matrix

$$W = \begin{bmatrix} W_1 \\ \vdots \\ W_m \end{bmatrix} = VY\mathcal{a} \qquad (D.7)$$

Equation (D. 6) now becomes

$$E\Phi_4\Phi_3\Phi_2\Phi_1 = \sum_\ell X_\ell W_\ell = \mathcal{J} Y W \qquad (D.8)$$

where

$$\mathcal{I} = (I, \underbrace{I, \ldots,}_{m \text{ terms}} I)$$

Substituting successively for W, V, U by using (D. 7), (D. 5), (D. 2) respectively, (D. 8) reduces to

$$E\,\Phi_4\Phi_3\Phi_2\Phi_1 = \mathcal{I}(YQ^T)^3 Y\mathcal{a}$$

which can be successively generalized to

$$E(\Phi_k\,\Phi_{k-1}\cdots\,\Phi_1) = \mathcal{I}(YQ^T)^{k-1}Y\mathcal{a}$$

APPENDIX E

SPECIALIZATION OF MARKOV CASE TO INDEPENDENT CASE

In Chap.III, Sec. 3. 2, it was shown that if

$$x_k = \Phi_k \cdots \Phi_1 x_0$$

with $\{\Phi_k\}$ an independent, identically distributed, matrix sequence, then $Ex_{k[i]} = (E\Phi_{[i]})^k x_{0[i]}$. In this appendix, this result is derived as a special case of Eq. (4. 8):

$$Ex_{k[i]} = \mathcal{I}_i(Y_iQ_i^T)^{k-1}Y_i\mathcal{a}_i x_{0[i]}$$

which is the result when $\{\Phi_k\}$ is a Markov chain.

If $\{\Phi_k\}$ is an independent process, then the Markov matrix has identical rows: $p_{ij} = p_j = P[\Phi_k = X_j]$ for all k. Hence $Y_iQ_i^T$ has the form

$$Y_iQ_i^T = \begin{bmatrix} p_1X_{1[i]} & p_1X_{1[i]} & \cdots & p_1X_{1[i]} \\ \vdots & \vdots & & \vdots \\ p_mX_{m[i]} & p_mX_{m[i]} & \cdots & p_mX_{m[i]} \end{bmatrix}$$

Also $a_i = p_i$, whence

$$Y_i\,a_i = \begin{bmatrix} p_1X_{1[i]} \\ \vdots \\ p_mX_{m[i]} \end{bmatrix}$$

Therefore,

$$Y_iQ_i^TY_i\,a_i = \begin{bmatrix} p_1X_{1[i]}E\,\Phi_{[i]} \\ \vdots \\ p_mX_{m[i]}E\,\Phi_{[i]} \end{bmatrix}$$

and

$$(Y_iQ_i^T)^2Y_i\,a_i = \begin{bmatrix} p_1X_{1[i]}\ (E\,\Phi_{[i]})^2 \\ \vdots \\ p_mX_{m[i]}\,(E\,\Phi_{[i]})^2 \end{bmatrix}$$

and, in general,

$$(Y_iQ_i^T)^{k-1}Y_i\,a_i = \begin{bmatrix} p_1X_{k[i]}\ (E\,\Phi_{[i]})^{k-1} \\ \vdots \\ p_mX_{m[i]}\,(E\,\Phi_{[i]})^{k-1} \end{bmatrix}$$

where

$$E\,\Phi_{[i]} = \sum_j p_j X_{j[i]}$$

Equation (4.8) now reduces to

$$Ex_{k[i]} = (E\,\Phi_{[i]})^k x_{0[i]}$$

which is the desired result.

APPENDIX F

COMPUTATION OF $\lambda(Y_i Q_i^T)$ FOR CYCLIC CHAINS

Suppose that the Markov chain is ergodic; that is, there is only one ergodic set and the transient set is empty. Suppose further that the ergodic set is cyclic, of period d. The modes are assumed to be suitably ordered so that the transition matrix E has the form (4. 4):

$$E \doteq \begin{bmatrix} 0 & C_1 & 0 & \cdots & 0 \\ 0 & 0 & C_2 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & C_{d-1} \\ C_d & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Hence the Y matrix has the form

$$Y_i = X_i^{(1)} \oplus \cdots \oplus X_i^{(d)}$$

where $X_i^{(j)}$ is the direct sum of the matrices $X_{\ell[i]}$ corresponding to the modes $\ell \in \mathcal{C}_j$.

$$Y_i Q_i^T = Y_i(E^T \otimes I_{n^i})$$

can now be written as

$$Y_i Q_i^T = \begin{bmatrix} 0 & 0 & \cdots & 0 & F_d \\ F_1 & 0 & \cdots & 0 & 0 \\ 0 & F_2 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & F_{d-1} & 0 \end{bmatrix}$$

where

$$F_j = X_i^{(j+1)} (C_j^T \otimes I_{n^i}) , \qquad j = 1, \ldots, d,$$

provided $X_i^{(d+1)}$ is interpreted as $X_i^{(1)}$.

Raising $Y_i Q_i^T$ to the dth power,

$$(Y_i Q_i^T)^d = F_1 F_d F_{d-1} \cdots F_2 \oplus F_2 F_1 F_d F_{d-1} \cdots F_3 \oplus \cdots \oplus F_d F_{d-1} \cdots F_2 F_1 .$$

All of the cyclic products of the F_j's on the right-hand side of the above equation are square, but not necessarily of the same order. Moreover, by Lemma 1 proven below, all of the cyclic products have the same eigenvalues, except possibly for the zero eigenvalue. Hence the nonzero eigenvalues of $Y_i Q_i^T$ are the nonzero eigenvalues of any of the cyclic products. Since only the nonzero eigenvalues of $Y_i Q_i^T$ are of interest, a convenient way to compute them is to compute the eigenvalues of the lowest order cyclic product.

LEMMA 1: Let A_j, $j = 1, \ldots, k$, be matrices of dimension $n_j \times n_{j+1}$ with $n_1 = n_{k+1}$. Then all the cyclic products $A_1 A_2 \ldots A_k$,

$A_2A_3 \ldots A_kA_1, \ldots, A_kA_1 \ldots A_{k-2}A_{k-1}$ have the same eigenvalues, except possibly for the zero eigenvalue.

The proof follows from

LEMMA 2: Let A be an r x s matrix and let B be an s x r matrix. Assume, for convenience, that $r \geq s$. Then,

(a) AB and BA are of dimension r x r and s x s respectively;

(b) zero is an eigenvalue of AB of multiplicity at least (r-s);

(c) every eigenvalue of BA is an eigenvalue of AB.

PROOF: (a) Follows from the definition of matrix multiplication.

(b) Let N be the r x (r-s) matrix all of whose elements are zero. Define $\bar{A}$, $\bar{B}$ to be the matrices

$$\bar{A} = [A, N] \qquad \bar{B} = \begin{bmatrix} B \\ N^T \end{bmatrix}$$

Then

$$\bar{A}\,\bar{B} = AB$$

$$\bar{B}\,\bar{A} = \begin{bmatrix} BA & 0 \\ 0 & 0 \end{bmatrix}$$

By inspection, zero is an eigenvalue, of multiplicity (r-s), of $\bar{B}\,\bar{A}$. But the products $\bar{A}\,\bar{B}$, $\bar{B}\,\bar{A}$ have the same eigenvalues since $\bar{A}$ and $\bar{B}$ are square matrices. Therefore, zero is an eigenvalue, of multiplicity (r-s), of AB ($= \bar{A}\,\bar{B}$), and

(c) every eigenvalue of BA is an eigenvalue of AB.

REFERENCES

1. J. C. Samuels and A. C. Eringen, "On Stochastic Linear Systems," Division of Engineering Sciences, Purdue University, Lafayette, Indiana, Technical Report No. 11, ONR, August 1957.

2. R. Bellman and R. Kalaba, "Functional Equations in Adaptive Processes and Random Transmission," IRE Trans., Vol. IT-5, Special Supplement, pp. 271-282, May 1959.

3. A. R. Bergen, "Stability of Systems with Randomly Time-Varying Parameters," IRE Trans., Vol. AC-5, No. 4, pp. 265-269, Sept. 1960.

4. B. H. Bharucha, "Stability of Random Linear Systems with Markov Parameters, I," Electronics Research Laboratory, University of California, Berkeley, Series No. 60, Issue No. 289, June 29, 1960.

5. E. H. Kerner, "The Band Structure of Mixed Linear Lattices," Proc. Phys. Soc., Sec. A, Vol. 69, Part 3, pp. 234-244, March 1956.

6. R. E. Kalman, "Analysis and Synthesis of Linear Systems Operating on Randomly Sampled Data," Electronics Research Laboratories, Columbia University, August 1957.

7. A. Rosenbloom, "Analysis of Linear Systems with Randomly Time-Varying Parameters," Proc. of the Symposium on Information Networks, Vol. III, pp. 144-154, April 1954, Polytechnic Institute of Brooklyn.

8. V. I. Tikhonov, "Fluctuation Action in the Simplest Parametric Systems," Automation and Remote Control, Vol. 18, No. 8, pp. 705-711, August 1958.

9. J. C. Samuels, "On the Mean Square Stability of Random Linear Systems," IRE Trans., Vol. IT-5, Special Supplement, pp. 248-259, May 1959.

10. A. R. Bergen, "Random Linear Systems: A Special Case," AIEE Paper No. 61-72, presented at the Winter General Meeting, New York, Jan.-Feb. 1961; to be published in Trans. AIEE, Applications and Industry.

11. L. A. Zadeh, "Frequency Analysis of Variable Networks," Proc. IRE, Vol. 38, No. 3, pp. 291-299, March 1950.

12. L. A. Zadeh, "On a Class of Stochastic Operators," J. of Math. and Physics, Vol. XXXII, No. 1, pp. 48-53, April 1953.

13. J. E. Bertram and P. E. Sarachik, "Stability of Circuits with Randomly Time-Varying Parameters," IRE Trans., Vol. IT-5, Special Supplement, pp. 260-270, May 1959.

14. R Bellman, "Limit Theorems for Non-Commutative Operations, I," Duke Math. J., Vol. 21, No. 3, pp. 491-500, September 1954.

15. H. Furstenberg and H. Kesten, "Products of Random Matrices," Ann. Math. Stat., Vol. 31, No. 2, pp. 456-469, June 1960.

16. L. Cesari, Asymptotic Behavior and Stability Problems in Ordinary Differential Equations, Springer-Verlag, Berlin, 1959; pp. 10-11.

17. H. Antosiewicz, "A Survey of Lyapunov's Second Method," in S. Lefschetz (ed.), Contributions to the Theory of Nonlinear Oscillations, Vol. IV, Princeton University Press, Princeton, N. J., 1958., pp. 141-143.

18. J. Massera, "Contributions to Stability Theory," Annals of Math., Vol. 64, No. 1, pp. 182-206, July 1956.

19. R. Kalman and J. Bertram, "Control System Analysis and Design via the 'Second Method' of Lyapunov, I and II," J. of Basic Engineering, Trans. ASME, Series D, Vol. 82, No. 2, pp. 371-400, June 1960.

20. A Householder, "The Approximate Solution of Matrix Problems," J. Assoc. Comp. Machinery, Vol. V, No. 3, July 1958, pp. 205-243.

21. M. Loéve, Probability Theory, 2nd Ed., Van Nostrand, New York, 1960; Chap. III.

22. C. C MacDuffee, The Theory of Matrices, Chelsea Publishing Co., New York, 1956.

23. H. Cramer, Mathematical Methods of Statistics, Princeton University Press, Princeton, N. J., 1946.

24. B. Friedman, Principles and Techniques of Applied Mathematics, John Wiley and Sons, New York, 1956.

25. M. Marden, The Geometry of the Zeros of a Polynomial in a Complex Variable, American Mathematical Society publication, New York, 1949.

26. E I. Jury and B. H. Bharucha, "Notes on the Stability Criterion for Linear Discrete Systems," letter to the Editor, IRE Trans., Vol. AC-6, No. 1, pp. 88-90, February 1961.

27. E. I. Jury, "Additions to Stability Criterion for Linear Discrete Systems," to be published as a letter to the Editor, IRE Trans. on Automatic Control.

28. E. I. Jury, "A Simplified Stability Criterion for Linear Discrete Systems," to be published as a technical report, Electronics Research Laboratory, University of California, Berkeley, Series No. 60.

29. A. T. Bharucha-Reid, Elements of the Theory of Markov Processes and Their Applications, McGraw-Hill, New York, 1960.

30. J. L. Doob, Stochastic Processes, John Wiley and Sons, New York, 1953.

31 F. R. Gantmacher, The Theory of Matrices, Vol. 2, Chelsea Publishing Co., New York, 1959.

32. J. G. Kemeny and J. L. Snell, Finite Markov Chains, Van Nostrand, New York, 1960.

33. D. L. Kuver, "On Some Criteria for the Boundedness of Solutions of Systems of Differential Equations," (in Russian) Akademia Nauk SSSR, Doklady, Nov. Ser., Vol. 69, No. 5, pp. 603-606, 1949.

34. P. R. Halmos, Measure Theory, Van Nostrand, New York, 1950; pp. 88-89.

35. E. A. Coddington and N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill, New York, 1955.

36. N. A. Naimark, Normed Rings, translated from the first Russian edition by Leo F. Boron, P. Noordhoff N. V., Groningen, The Netherlands, 1959, pp. 49-51.

DISTRIBUTION LIST

Contract AF 19(604)-5466

Code	ORGANIZATION	NO. COPIES
AF 5	AFMTC (AFMTC Tech. Library - MU-135 Patrick AFB, Florida	1
AF 18	AUL Maxwell AFB, Alabama	1
AF 43	ASD (ASAPRD-Dist) Wright Patterson AFB, Ohio	1
AF 91	AFOSR (SRGL) Washington 25, D. C.	1
AF 124	RADC (RAYLD) Griffiss AFB, New York Attn: Documents Library	1
AF 139	AF Missile Development Center (MDGRT) Holloman AFB, New Mexico	1
AF 318	ARL (Technical Library) Building 450 Wright-Patterson AFB, Ohio	1
Ar 5	Commanding General USASRDL Fort Monmouth, New Jersey Attn: Tech. Doc. Ctr. SIGRA/SL-ADT	1
Ar 9	Department of the Army Office of the Chief Signal Officer Washington 25, D. C. Attn: SIGRD-4a-2	1
Ar 50	Commanding Officer Diamond Ordnance Fuze Laboratories Washington 25, D. C. Attn: ORDTL-012	1
Ar 67	Army Rocket and Guided Missile Agency Redstone Arsenal, Alabama Attn: ORDXR-OTL, Technical Library	1
G 2	ASTIA (TIPAA) Arlington Hall Station Arlington 12, Virginia	10
G 68	National Aeronautics and Space Agency 1520 H Street, N. W. Washington 25, D. C. Attn: Library	1
G 109	Director Langley Research Center National Aeronautics and Space Administration Langley Field, Virginia	1
N 140	Office of Naval Research Code 437, Information Systems Branch Washington 25, D. C. Attn: Marshall C. Yovits	1
M 6	AFCRL, OAR (CRIPA - Stop 39) L. G. Hanscom Field Bedford, Massachusetts	10
M 78	AFCRL, OAR (CRT, Dr. A. M. Gerlach) L. G. Hanscom Field Bedford, Massachusetts	1
N 1	Director, Avionics Division (AV) Bureau of Aeronautics Department of the Navy Washington 25, D. C.	2
N 29	Director (Code 2027) U.S. Naval Research Laboratory Washington 25, D. C.	2
I 292	Director, USAF Project RAND The Rand Corporation 1700 Main Street Santa Monica, California Thru: A. F. Liaison Office	1
AF 127	Boston Office - Patents& Royalties Division (Hq AFLC) Building 133, 424 Trapelo Road Waltham 54, Massachusetts	1
AF 253	Technical Information Office European Office, Aerospace Research Shell Building, 47 Cantersteen Brussels, Belgium	1
Ar 107	U.S. Army Aviation Human Research Unit U.S. Continental Army Command P. O. Box 438 Fort Rucker, Alabama Attn: Maj. Arne H. Eliasson	1
G 8	Library Boulder Laboratories National Bureau of Standards Boulder, Colorado	2
M 63	Institute of the Aerospace Sciences, Inc. 2 East 64th Street New York 21, New York Attn: Librarian	1
N 73	Office of Naval Research Branch Office, London Navy 100, Box 39 F. P. O. New York, New York	10
U 32	Massachusetts Institute of Technology Research Laboratory of Electronics Building 26, Room 327 Cambridge 39, Massachusetts Attn: John H. Hewitt	1
G 73	National Science Foundation Engineering Sciences Program 1951 Constitution Avenue Washington 25, D. C.	1
U 431	Alderman Library University of Virginia Charlottesville, Virginia	1
I 881	American Systems Incorporated 1625 East 126th Street Hawthorne, California Attn: M. D. Adcock, Director Electromagnetic Systems Div.	1
M 62	Major General Casemiro Montenegro Filho, Director Centro Tecnico da Aeronautica (CTA) San Jose dos Compos Sao Paulo, Brazil	1
U 413	The Ohio State University Antenna Laboratory 2024 Neil Avenue Columbus 10, Ohio Attn: Security Officer	1
U 414	Princeton University School of Engineering Princeton, New Jersey Attn: Miss Karen Takle, Librarian The Engineering Library	1
U 415	Technological University Kanaalweg 2B Delft, Netherlands Attn: Dr. Pieter Eykhoff Department of Electrical Engineering	1
AF 179	AFOSR (Harold Wooster) Director, Research Information Office Washington 25, D. C.	1
AF 237	WADD (WCLJR) Wright-Patterson AFB, Ohio Attn: Mr. R. M. Lemmen	1
AF 238	RADC (RCWID) Griffiss AFB, New York Attn: Capt. B. J. Long	1
AF 239	Atomic Energy Commission (John R. Pasta) Division of Research Washington 25, D. C.	1
AF 255	496L SPO (Space Track) (ESSIS, Stop 2) 424 Trapelo Rd Waltham 54, Massachusetts	1
G 72	Roger Moulton National Security Agency Ft. George G. Meade, Maryland	1
Remaining copies to:	AFCRL, Office of Aerospace Research (CRRB) L. G. Hanscom Field Bedford, Massachusetts Attn: Hans H. Zschirnt	5